

**PAIN SIGNALING MOLECULES**

5 This application claims priority under 35 U.S.C. §120 to U.S. Patent Application No. 09/704,707, filed November 3, 2000 and under 35 U.S.C. §119(e) to U.S. Provisional Applications 60/202,027, filed May 4, 2000, 60/222,344, filed August 1, 2000, and 60/ , filed April 19, 2001, which are herein incorporated by reference in their entirety.

10 Background of the Invention

Field of the Invention

The invention relates generally to novel genes expressed in normal but not Neurogenin-1-deficient animals. The invention relates specifically to a novel family of G protein-coupled receptors and a novel family of two-transmembrane segment proteins  
15 that are expressed in dorsal root ganglia, and a method of screening for genes specifically expressed in nociceptive sensory neurons.

Description of the Related Art

20 The treatment of acute and chronic intractable pain is a major target of drug development in the pharmaceutical industry. Pain sensation is mediated by primary sensory neurons in the dorsal root ganglia (DRG), which project peripherally to the skin and centrally to the spinal cord. These neurons express signaling molecules, such as receptors, ion channels and neuropeptides, which are involved in pain sensation. One  
25 example is the so-called Vanilloid Receptor-1 (VR-1), which is activated by capsaicin (chili pepper) as well as by heat and acid. Such pain signaling molecules may also influence pain sensation indirectly by acting as positive or negative modulators of the sensory pathway. Searching for drugs that interact with such signaling molecules, for example as receptor agonists or antagonists, is an important approach to the discovery  
30 of new therapeutics for the treatment of pain. New candidate signaling molecules expressed by pain-sensing ("nociceptive") sensory neurons are therefore highly

desirable targets for new drug screening and drug discovery efforts. The present inventors have previously identified a novel family of basic helix-loop-helix (bHLH) transcription factors, called the *Neurogenins* (*Ngns*), which are essential for the development of sensory neurons in the DRG. Different *Ngns* are required for the development of different subsets of sensory neurons. In particular, *Ngn1* is necessary for the development of most if not all nociceptive sensory neurons. In *Ngn1*<sup>-/-</sup> mutant mice, although DRG are still present, they are reduced in size and the majority of nociceptive neurons, identified by expression of markers such as trkA and VR-1, are missing (Ma *et al.* Genes&Develop, 13: 1717-1728, (1999)). These results suggested that the isolation of genes expressed in wild-type (normal) but not *Ngn1*<sup>-/-</sup> DRG might lead to the identification of novel drug target molecules expressed in differentiating or mature nociceptive sensory neurons.

While pain is usually a natural consequence of tissue injury, as the healing process commences the pain and tenderness associated with the injury resolve. However, some individuals experience pain without an obvious injury or suffer protracted pain after an initial insult. In addition, chronic or intractable pain may occur in association with certain illnesses, such as, for example, bone degenerative diseases, terminal cancer, AIDS, and Reflex sympathetic dystrophy (RSD). Such patients may be unable to receive relief with currently-available pain-relieving (anti-nociceptive) drugs, such as opioid compounds, e.g. morphine, due to problems such as dependence and tolerance. Therefore, there is a great need for novel therapeutic agents for the treatment of pain, in particular chronic pain.

#### Summary of the Invention

The present inventors have carried out a screen for genes expressed in wild-type but not *Ngn1*<sup>-/-</sup> DRG using positive selection-based differential hybridization. This screen has identified both known signaling molecules involved in nociceptive neuron function, such as VR-1, and novel signaling molecules that are highly specifically expressed in nociceptive sensory neurons. The present invention therefore includes the discovery of new genes that are expressed in normal mice but not in *Ngn1* null mutant mice. One family of novel genes isolated from the screen encodes a receptor protein

with 7 transmembrane segments, *mrg*, a characteristic of G protein-coupled receptors. Subsequent staining experiments (see Fig. 2, 2A-D) confirmed that *mrg* genes were expressed specifically in subsets of nociceptive neurons in DRG. Another novel gene family isolated in this screen, *drg-12*, encodes a protein with two transmembrane segments.

In particular, the invention includes isolated nucleic acid molecules that encode a *mrg* protein selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25, 27, an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25, 27, an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule comprising SEQ ID NOS: 1, 3, 5, 7, 9, 11, 15, 17, 20, 22, 24 or 26 and an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27. Nucleic acid molecules of the invention also include those that encode a protein that is expressed in dorsal root ganglia and have at least about 60% nucleotide sequence identity, preferably at least about 70-75% sequence identity, more preferably at least about 80-85% sequence identity, and even more preferably at least about 90% sequence identity through the coding sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 15, 17, 20, 22, 24 or 26. Alternatively, nucleic acid molecules of the invention may encode a *mrg* protein that exhibits at least about 38% amino acid sequence identity with SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27.

The invention also includes isolated nucleic acid molecules that encode a *drg-12* protein selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of SEQ ID NOS: 14, 19 or 29 an isolated nucleic acid molecule that encodes a fragment of at least 6 amino acids of SEQ ID NOS: 14, 19 or 29, an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule comprising SEQ ID NO: 13 or 28 and an isolated nucleic acid molecule which hybridizes to the complement of a nucleic acid molecule that encodes the amino acid sequence of SEQ ID NOS: 14, 19 or 29. Nucleic acid molecules of the invention also include those that encode a *drg-12* protein that is expressed in dorsal root ganglia

and have at least about 60% nucleotide sequence identity, preferably at least about 70-75% sequence identity, more preferably at least about 80-85% sequence identity, and even more preferably at least about 90% sequence identity through the coding sequence of SEQ ID NO: 13 or 28. Alternatively, nucleic acid molecules of the invention may  
5 encode a *drg-12* protein that exhibits at least about 33% amino acid sequence identity with SEQ ID NOS: 14, 19 or 29.

The present invention also includes the nucleic acid molecules described above operably linked to one or more expression control elements, including vectors comprising the isolated nucleic acid molecules. The invention further includes host  
10 cells transformed to contain the nucleic acid molecules of the invention and methods for producing a protein comprising the step of culturing a host cell transformed with a nucleic acid molecule of the invention under conditions in which the protein is expressed.

The invention further provides an isolated Mrg polypeptide selected from the  
15 group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27, an isolated polypeptide comprising a functional fragment of at least 10 amino acids of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27, an isolated polypeptide comprising conservative amino acid substitutions of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27 and naturally  
20 occurring amino acid sequence variants of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27. Polypeptides of the invention also include polypeptides with an amino acid sequence having at least about 38%, 40%, 50%, 60%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25 or 27 more preferably at least about 80%, even more preferably at least about  
25 90%, and most preferably at least about 95% sequence identity with these sequences.

The invention further provides an isolated Drg-12 polypeptide selected from the group consisting of an isolated polypeptide comprising the amino acid sequence of SEQ ID NOS: 14, 19 or 29, an isolated polypeptide comprising a functional fragment of at least 10 amino acids of SEQ ID NOS: 14, 19 or 29 an isolated polypeptide comprising  
30 conservative amino acid substitutions of SEQ ID NOS: 14, 19 or 29 and naturally occurring amino acid sequence variants of SEQ ID NOS: 14, 19 or 29. Polypeptides of

the invention also include polypeptides with an amino acid sequence having at least about 33%, 35%, 40%, 50%, 60%, 70% or 75% amino acid sequence identity with the sequence set forth in SEQ ID NO: 14, 19 or 29, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% sequence identity with these sequences.

The invention further provides an isolated antibody that specifically binds to a polypeptide of the invention, including monoclonal and polyclonal antibodies.

The invention further provides methods of identifying an agent which modulates the expression of a nucleic acid encoding a protein of the invention, comprising the steps of: exposing cells which express the nucleic acid to the agent; and determining whether the agent modulates expression of such nucleic acid, thereby identifying an agent which modulates the expression of a nucleic acid encoding the protein.

The invention further provides methods of identifying an agent which modulates at least one activity of a protein of the invention, comprising the steps of: exposing cells which express the protein to the agent; and determining whether the agent modulates at least one activity of the protein, thereby identifying an agent which modulates at least one activity of the protein.

The invention further provides methods of identifying binding partners for a protein of the invention, comprising the steps of: exposing said protein to a potential binding partner; and determining if the potential binding partner binds to the protein, thereby identifying binding partners for the protein.

The present invention further provides methods of modulating the expression of a nucleic acid encoding a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein. The invention also provides methods of modulating at least one activity of a protein of the invention, comprising the step of: administering an effective amount of an agent which modulates at least one activity of the protein.

The present invention further includes non-human transgenic animals modified to contain the nucleic acid molecules of the invention or mutated nucleic acid molecules such that expression of the polypeptides of the invention is prevented.

The invention further provides methods of pain treatment, comprising the steps of: administering to a patient in need thereof a therapeutically effective amount of an agent that modulates the production or at least one activity of a polypeptide or nucleic acid of the invention.

5 In another aspect the invention provides a method of identifying candidate genes involved in nociception comprising the steps of: generating a first set of non-human animals that is *Ngn1*<sup>-/-</sup> and a second set of non-human animals that is wild-type for the *Ngn1* gene; isolating RNA from the dorsal root ganglia of the first and second set of animals; enriching for genes expressed in the DRG wild-type but not in the *Ngn1* mutant animals; and further characterizing and selecting for candidate genes using  
10 methods such as sequencing, degenerated RT-PCR and *in situ* hybridization.

#### Brief Description of the Drawings

15 Figure 1 shows the alignment of a homologous region of the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10 and 12, and also of two human members of the *mrg* family (SEQ ID NOS: 16 and 18).

Figure 1A indicates that *mrgs* define a Novel G protein-couple receptor Gene Family. Amino acid sequences of eight mouse full-length *mrg* genes were aligned using  
20 ClustalW. Identical residues in >50% of the predicted proteins are darkly shaded; conservative substitutions are highlighted in light gray. The approximate locations of predicted transmembrane domain 1-7 are indicated on top of the sequences as TM1-TM7. The predicted extracellular and cytoplasmic domains are indicated as E1-E7 and C1-C7 respectively.

25 The microscopy images of *in situ* hybridization in Figure 2 show the localization of antisense staining against a nucleotide of SEQ ID NO: 2 ("*mrg3*") and of SEQ ID NO: 4 ("*mrg4*") in transverse sections of dorsal root ganglia (DRG) from newborn wild type (WT) and *Neurogenin1* null mutant (*Ngn1*<sup>-/-</sup>) mice. White dashed lines outline the DRG and black dashed lines outline the spinal cord. Note that in the *Ngn1*<sup>-/-</sup> mutant,  
30 the size of the DRG is severely reduced due to the loss of nociceptive sensory neurons, identified using three other independent markers (*trkA*; VR-1 and SNS-TTXi (Ma et al.,

(1999)). *mrg3* is expressed in a subset of DRG in WT mice (A) but is absent in the *Ngn1*<sup>-/-</sup> DRG (B). *mrg4* is expressed in a smaller subset of DRG than that of *mrg3* (C). It is also absent in the *Ngn1*<sup>-/-</sup> DRG (D). The loss of *mrg*-expressing neurons in the *Ngn1*<sup>-/-</sup> DRG indicates that these neurons are likely to be nociceptive.

5           Figure 2A shows expression of *mrgs* in subsets of dorsal root ganglia (DRG) neurons. Frozen transverse sections of DRG from wild-type (a-i) and *ngn1*<sup>-/-</sup> (j) mutant new born mice were annealed with antisense digoxigenin RNA probes, and hybridization was visualized with an alkaliine phosphatase-conjugated antibody. Positive signals are shown as dark purple stainings. TrkA is expressed in a large portion of wild-type DRG neurons (a) but absent in *ngn1*<sup>-/-</sup> (data not shown). Each of the eight *mrg* genes (b-i) is expressed in a small subset of neurons in wild-type DRG in completely absent in *ngn1*<sup>-/-</sup> DRG (j and data not shown). Black dash line outlines the *ngn1*<sup>-/-</sup> mutant DRG.

15           Figure 2B shows that *mrgs* are expressed by TrkA<sup>+</sup> nociceptive neurons. Double labeling technique was used to colocalize TrkA (b,e) and *mrgs* (a,d) in DRG neurons. During the double labeling experiments frozen sections of wild-type DRG were undergone in situ hybridizations with either *mrg3* (a-c) or *mrg5* (d-f) fluorescein-labeled antisense RNA probes followed by anti-TrkA antibody immunostaining. The same two frames (a and b, d and e) were digitally superimposed to reveal the extent of colocalization (c, f). The white arrowheads indicate examples of double positive cells.

20           Figure 2C shows that *mrgs* and VR1 define two different populations of nociceptive neurons in DRG. The combination of in situ hybridizations with either *mrg3* or *mrg5* fluorescein-labeled antisense RNA probes and anti-VR1 antibody immunostaining demonstrated that neither *mrg3* (a-c) nor *mrg5* (d-f) were expressed by VR1-positive neurons. In the merged images (c,f), there are no colocalizations of VR1 with either *mrg3* or *mrg5*. The white arrowheads are pointed to *mrgs*-expressing but VR1-negative nociceptive neurons.

25           Figure 2D shows that *mrgs* are expressed by IB4<sup>+</sup> nociceptive neurons. Double labeling technique was used to colocalize IB4 (b,e) and *mrgs* (a,d) in DRG neurons. The expressions of *mrg3* and *mrg5* were visualized by in situ hybridization as described before. The same DRG sections were subsequently undergone through FITC-

conjugated lectin IB4 binding. In the merged images (c,f), there are extensive overlappings between mrgs and IB4 stainings (yellow neurons indicated by arrowheads).

Figure 3 compares the hydrophobicity plots predicting the transmembrane regions of the amino acid sequence of (A) *mrg3* (SEQ ID NO: 2); (B) human1 gene (SEQ ID NO: 15); and (C) human2 gene (SEQ ID NO: 17). More positive values indicate hydrophobicity.

Figure 4 compares the hydrophobicity plots predicting the transmembrane regions of the amino acid sequence of (A) mouse *drg12* (SEQ ID NO: 14); (B) human *drg12* (SEQ ID NO: 19)

Figure 5 compares the hydrophobicity plots predicting the transmembrane regions of the amino acid sequence of *mrg9* (SEQ ID NO: 21); *mrg10* (SEQ ID NO: 23); *mrg11* (SEQ ID NO: 25) and *mrg12* (SEQ ID NO: 27).

Figure 6A is an alignment of the amino acid sequences of MRGA1-A8, deduced from nucleotide sequences of cDNA and BAC clones from strain C57BL/6J mice. The predicted locations of the transmembrane (TM1-TM7), extracellular (E1-E4), and cytoplasmic (C1-C4) domains are indicated above the aligned sequences.

Figure 6B depicts a phylogenetic analysis of MRG family members identified from database searches. The protein sequences of all MRGs were aligned using CLUSTALW (Thompson et al. Nucleic Acids Res 22: 4673-80 (1994)). The dendrogram was generated with the PHYLUP software package using the Neighbor-Joining method and 1,000 bootstrap trials. The horizontal length of the branches is proportional to the number of amino acid changes. Vertical distances are arbitrary. Mouse (m)*Mrg* genes with retrotransposon sequences ~650 nt 3' of their stop codon are highlighted (L1). All genes that are predicted to encode pseudogenes are indicated with the psi ( $\Psi$ ) symbol.

Figure 6C shows the chromosomal organization of one mouse *Mrg* cluster deduced from analysis of overlapping BAC clones. The cluster contains four intact ORFs and three pseudogenes.

Figure 7A shows the distribution of nociceptive sensory neurons in a postnatal day 0 (P0) DRG as revealed by expression of the NGF receptor *trkA*. This population is



selectively eliminated in *Ngn1*<sup>-/-</sup> mutants (Ma et al. Genes & Dev. 13: 1717-1728 (1999)).

Figure 7B shows in situ hybridization with cRNA probes detecting *MrgA1*. *MrgA1* is expressed in a pattern similar to that of *trkA*<sup>+</sup> neurons on an adjacent section shown in Figure 7A.

Figure 7C shows in situ hybridization with cRNA probes detecting *MrgA2*-*MrgA8*.

Figure 7J shows that *MrgA1* expression is eliminated in *Ngn1*<sup>-/-</sup> mice, as is expression of other *MrgA* genes (not shown). Remaining DRG neurons are present in the area delimited by the dotted line, and can be visualized by expression of generic neuronal markers.

Figure 8 shows that expression of *MrgAs* is restricted to non-peptidergic nociceptors that project to inner lamina II. Shown are confocal microscopic images of in situ hybridizations using the *Mrg* probes indicated, combined with fluorescent antibody detection of *trkA* (A-D), substance P (I-L), CGRP (M-P), VR1 (Q-T) or staining with fluorescent isolectin IB4 (IB4; E-H). *MrgA*<sup>+</sup> or *MrgD*<sup>+</sup> cells co-express *trkA* and IB4 (A-H, arrowheads), but most do not express subP, CGRP or VR1 (I-T, arrowheads; arrows in I, M indicate a minor subset of *MrgA1*<sup>+</sup> neurons that co-express SubP and CGRP).

Figure 9 is a schematic illustration of the restriction of *MrgA* (and *MrgD*) expression to non-peptidergic, IB4<sup>+</sup>, VR1<sup>-</sup> sensory neurons that project to lamina III (Snider and McMahon Neuron 20: 629-32 (1998)). Post-synaptic neurons in lamina III express PKCγ.

Figure 10 shows that individual sensory neurons co-express multiple *MrgAs*. (A-C) double label in situ hybridization with *MrgA1* (A) and *A3* (B). (D-F) double labeling with *MrgA1* (D) and *MrgA4* (E). In both cases, cells expressing *MrgA3* or *A4* are a subset of those expressing *MrgA1* (C, F, arrowheads). Arrows in (F) indicate intranuclear dots of *MrgA4* expression which may represent sites of transcription. (G-I) Double label in situ with *MrgA1* and *MrgD*. Some overlap between the two populations is seen (I, arrowhead), while most cells express one receptor but not the other (I, arrows). Approximately 15% of cells expressing either *MrgA1* or *MrgD* co-

express both genes. Vertical bars to the right of panels (C, F, I) represent a z-series viewed along the y-axis, horizontal bars below the panels a z-series viewed along the x-axis. (J, K) comparison of in situ hybridization signals obtained using a single *MrgA* probe (J) and a mixture of 7 *MrgA* probes (K). Approximately 1% of neurons were labeled by the *MrgA4* probe, while ~4.5% were labeled by the mixed probe. The sum of the percentage of neurons labeled by the individual *MrgA2-8* probes is ~6.6%, suggesting that there is partial overlap within this population. (L) Venn diagram illustrating combinations of gene expression revealed by in situ analysis. The drawing is a conservative estimate of the number of subsets, since we do not yet know, for example, whether *MrgAs2-8* partially overlap with *MrgD*. The sizes of the circles are not proportional.

Figure 11 shows elevated intracellular free  $\text{Ca}^{++}$  elicited by FLRF in HEK cells expressing MRGA1. (A, B) and (E, F) illustrate Fura-2 fluorescence at 340 nm (A, E) and 380 nm (B, F) in HEK- $\text{G}\alpha_{15}$  cells expressing an MRGA1-GFP fusion protein (A-D) or GFP alone (E-H). The images were taken 2 minutes after the addition of 1  $\mu\text{M}$  of FLRFamide. The peri-nuclear, punctate distribution of MRGA1-GFP revealed by intrinsic GFP fluorescence (D, arrowheads) is characteristic of the ER-Golgi network, indicating membrane integration and intracellular transport of the receptors. In contrast, the control GFP protein is cytoplasmic (H). The intracellular  $\text{Ca}^{2+}$  ( $[\text{Ca}^{2+}]_i$ ) release was determined from the FURA-2 340nm/380nm emission ratio (C, G). Note that MRGA1-expressing cells (but not surrounding untransfected cells) show an elevated ratio of Fura-2 fluorescence at 340/380 nm (C, arrowheads), indicating an increase in  $[\text{Ca}^{2+}]_i$ . In contrast, no such elevation is observed in control GFP-expressing cells (G). The elevated 340/380 fluorescence seen in MRGA1-expressing cells was dependent on the addition of ligand (not shown).

Figure 12A shows activation of MRGA receptors expressed in heterologous cells by neuropeptide ligands. HEK- $\text{G}\alpha_{15}$  cells (Offermanns and Simon. *J Biol Chem* 270: 15175-80 (1995)) expressing MRGA1 were tested with the indicated ligands at a concentration of 1  $\mu\text{M}$ . The data indicate the mean percentages of GFP-positive (i.e., transfected) cells showing calcium responses. None of the agonists indicated showed any responses through endogenous receptors in untransfected cells. Note that the

RFamide neuropeptides FMRF, FLRF and NPFF, as well as NPY, ACTH, CGRP-I and -II and somatostatin (SST) produced the strongest responses.

Figure 12B shows the ligand selectivity of MRGA1 expressed in HEK cells lacking  $G\alpha_{15}$ . The cells were exposed to ligands at a concentration of 1  $\mu$ M as in (A).

Figure 12C shows the ligand selectivity of MRGA4. The data presented in Figures 12B and 12C indicate that the responses to the most effective ligands do not depend on the presence of  $G\alpha_{15}$ . Note that MRGA1-expressing cells respond to FLRF and NPFF but not to NPAF, while conversely MRGA4-expressing cells respond to NPAF but not NPFF or FLRF

Figure 12D shows dose-response curves for MRGA1 expressed in HEK- $G\alpha_{15}$  cells to selected RFamide neuropeptides. Each data point represents the mean  $\pm$ S.E.M. of at least 3 independent determinations; at least 20 GFP<sup>+</sup> cells were analyzed for each determination. Responses at each ligand concentration were normalized to the maximal response subsequently shown by the same cells to a 5  $\mu$ M concentration of FLRF. MRGA1 (D) shows highest sensitivity to FLRF (squares,  $EC_{50} \approx 20$  nM) and lower sensitivity to NPFF (circles,  $EC_{50} \approx 200$  nM).

Figure 12E shows dose-response curves for MRGA4 expressed in HEK- $G\alpha_{15}$  cells to selected RFamide neuropeptides. Each data point represents the mean  $\pm$ S.E.M. of at least 3 independent determinations; at least 20 GFP<sup>+</sup> cells were analyzed for each determination. Responses at each ligand concentration were normalized to the maximal response subsequently shown by the same cells to a 5  $\mu$ M concentration of NPAF. MRGA4 is preferentially activated by NPAF (triangles,  $EC_{50} \approx 60$  nM).

Figure 12F shows dose-response curves for MAS1 expressed in HEK- $G\alpha_{15}$  cells to selected RFamide neuropeptides. Each data point represents the mean  $\pm$ S.E.M. of at least 3 independent determinations; at least 20 GFP<sup>+</sup> cells were analyzed for each determination. Responses at each ligand concentration were normalized to the maximal response subsequently shown by the same cells to a 5  $\mu$ M concentration of NPFF. MAS1, like MRGA1, is activated by NPFF with similar efficacy ( $EC_{50} \approx 400$  nM), but is not as well activated by FLRF (squares).

Figure 13 depicts the expression pattern of mMrgB1 in a sagittal section of a newborn mouse. The staining pattern indicates that the mMrgB1 gene is expressed in

the scattered cells in the epidermal layer of the skin, in the spleen and in the submandibular gland.

Figure 14 is a higher magnification of the mMrgB1 expression in the spleen and skin depicted in Figure 13.

5                   Figure 15 shows the expression of mMrgD in adult dorsal root ganglia.

## Detailed Description of the Preferred Embodiment

### **I. General Description**

10                   As described above, the present invention is based on the discovery of new genes that are expressed in the DRG of normal mice but not in *Ngn1* null mutant mice. One of the novel gene families isolated from the screen encodes a receptor protein with 7 transmembrane segments, a characteristic of G protein-coupled receptors. This novel 7 transmembrane receptor is most closely related to the oncogene *mas*, and therefore  
15                   was provisionally named *mas-related gene-3* (*mrg3*). *mrg3* is now known as *MrgA1*, and the terms are used interchangeably herein. Almost 50 members of the *Mas-related gene* (*Mrg*) family have been identified, many of which are specifically expressed in non-peptidergic nociceptors. Large families of G protein-coupled receptors are also expressed in other classes of sensory neurons, such as olfactory and gustatory neurons.

20                   The murine *Mrg* family of GPCRs contains three major subfamilies (*MrgA*, *B* and *C*), each consisting of more than 10 highly duplicated genes, as well as several single-copy genes such as *Mas1*, *Rta*, *MrgD* and *MrgE* (Figure 6B). The *MrgA* subfamily consists of at least twenty members in mice: *MrgA1* (SEQ ID NO: 2); *MrgA2* (SEQ ID NO: 4); *MrgA3* (SEQ ID NO: 6); *MrgA4* (SEQ ID NO: 11); *MrgA5* (SEQ ID NO: 21); *MrgA6* (SEQ ID NO: 23); *MrgA7* (SEQ ID NO: 25); *MrgA8* (SEQ  
25                   ID NO: 27); *MrgA9* (SEQ ID NO: 53); *MrgA10* (SEQ ID NO: 55); *MrgA11* (SEQ ID NO: 57); *MrgA12* (SEQ ID NO: 59); *MrgA13* (SEQ ID NO: 61); *MrgA14* (SEQ ID NO: 63); *MrgA15* (SEQ ID NO: 65); *MrgA16* (SEQ ID NO: 67); *MrgA17* (SEQ ID NO: 69); *MrgA18* (SEQ ID NO: 71); *MrgA19* (SEQ ID NO: 73); *MrgA20* (SEQ ID  
30                   NO: 75). Four human sequences that are most closely related to the *MrgA* subfamily

have also been identified: MrgX1 (SEQ ID NO: 16); MrgX2 (SEQ ID NO: 18); MrgX3 (SEQ ID NO: 31); and MrgX4 (SEQ ID NO: 33).

The MrgB subfamily consists of at least twelve members in mice: MrgB1 (SEQ ID NO: 39); MrgB2 (SEQ ID NO: 41); MrgB3 (SEQ ID NO: 43); MrgB4 (SEQ ID NO: 45); MrgB5 (SEQ ID NO: 47); MrgB6 (SEQ ID NO: 77); MrgB7 (SEQ ID NO: 79); MrgB8 (SEQ ID NO: 81); MrgB9 (SEQ ID NO: 83); MrgB10 (SEQ ID NO: 85); MrgB11 (SEQ ID NO: 87); and MrgB12 (SEQ ID NO: 89).

Ten members of the MrgC subfamily have been identified in mice: MrgC1 (SEQ ID NO: 91); MrgC2 (SEQ ID NO: 93); MrgC3 (SEQ ID NO: 95); MrgC4 (SEQ ID NO: 97); MrgC5 (SEQ ID NO: 99); MrgC6 (SEQ ID NO: 101); MrgC7 (SEQ ID NO: 103); MrgC8 (SEQ ID NO: 105); MrgC9 (SEQ ID NO: 107); and MrgC10 (SEQ ID NO: 109).

A single member of the MrgD subfamily has been identified in mice, mMrgD (SEQ ID NO: 49) and its ortholog identified in humans, hMrgD (SEQ ID NO: 35). Similarly, a single member of the MrgE subfamily has been identified in mice, mMrgE (SEQ ID NO: 51) and humans, hMrgE (SEQ ID NO: 37).

As is the case in other GPCR subfamilies, a number of the *Mrgs* appear to be pseudogenes, including all members of the *MrgC* subfamily. The presence of L1 retrotransposon elements near several *Mrg* genes raises the possibility that pseudogene expansion may have been driven by L1-mediated transduction (Goodier et al. Hum Mol Genet 9: 653-7 (2000)).

In contrast to the murine *MrgA* and *B* subfamilies, which together contain almost 40 intact coding sequences, only four intact human *MrgX* sequences were identified. The remaining 10 human *Mrg* sequences appear to be pseudogenes. Inclusion of other related receptors such as *hMrgD* and *hMas1* brings the total number of intact human coding sequences in this family to nine (Figure 6B).

Prior to the present invention, the primary nociceptive sensory neurons were thought not to specifically discriminate among different chemical stimuli, but rather to detect noxious stimuli of various modalities by virtue of broadly tuned receptors such as VR1 (Tominaga et al. Neuron 21: 531-43 (1998)). The expression of *Mrgs* reveals an unexpected degree of molecular diversification among nociceptive sensory neurons.

Approximately 13-14% of sensory neurons express *MrgA1*, while 17-18% express *MrgD* and the overlap between these two populations is only 15%. The *MrgA1*<sup>+</sup> population seems to include most or all neurons expressing *MrgA2-8*. However, these latter *MrgA* genes are not all expressed in the same neurons. Thus the 8 *MrgA* genes and *MrgD* define at least 6 different neuronal subpopulations, and the remaining 16 *MrgA* genes add even greater diversity.

It is striking that both *MrgA* and *D* are expressed in IB4<sup>+</sup>, VR1<sup>-</sup> sensory neurons. IB4<sup>+</sup> neurons are known to project to lamina III (Snider and McMahon Neuron 20: 629-32 (1998)), which has been implicated in chronic pain, such as that accompanying nerve injury (Malmberg et al. Science 278: 279-83 (1997)). VR1 is activated both by thermal stimuli and chemical stimuli such as capsaicin (Caterina et al. Nature 389: 816-824 (1997); Tominaga et al. Neuron 21: 531-43 (1998)), but VR1<sup>+</sup> neurons are dispensable for the detection of noxious mechanical stimuli (Caterina et al. Science 288: 306-13 (2000)). This indicates that one of the functions of *MrgA*<sup>+</sup> neurons is the detection of noxious mechanical stimuli accompanying neuropathic or inflammatory pain.

The existence of a family of putative G protein-coupled receptors specifically expressed in nociceptive sensory neurons suggests that these molecules are primary mediators or modulators of pain sensation. It is therefore of great interest to identify ligands, both endogenous and synthetic, that modulate the activity of these receptors, for the management of chronic intractable pain. Indeed, ligand screens in heterologous cell expression systems indicate that these receptors can interact with RF-amide neuropeptides of which the prototypic member is the molluscan cardioexcitatory peptide FMRF-amide (Price and Greenberg Science 197: 670-671 (1977)). Mammalian RF-amide peptides include NPFF and NPAF, which are derived from a common pro-peptide precursor expressed in neurons of laminae I and II of the dorsal spinal cord (Vilim et al. Mol Pharmacol 55: 804-11 (1999)). The expression of this neuropeptide FF precursor in the synaptic termination zone of *Mrg*-expressing sensory neurons, the ability of NPAF and NPFF to activate these receptors in functional assays, and the presence of binding sites for such peptides on primary sensory afferents in the dorsal horn (Gouarderes et al. Synapse 35: 45-52 (2000)), together indicate that these neuropeptides are ligands for *Mrg* receptors *in vivo*. As intrathecal injection of

NPFF/NPAF peptides produces long-lasting antinociceptive effects in several chronic pain models (reviewed in Panula et al. Brain Res 848: 191-6 (1999)), including neuropathic pain (Xu et al. Peptides 20: 1071-7 (1999)), these data further indicate that *Mrgs* are directly involved in the modulation of pain.

5           One possibility for the extent of diversity among *Mrgs* expressed by murine nociceptors is that different *Mrgs* are expressed by sensory neurons that innervate different peripheral targets, such as gut, skin, hair follicles, blood vessels, bones and muscle. These targets may secrete different ligands for different *Mrgs*. Another possibility is that neurons expressing different *Mrgs* respond to a common modulator of peripheral nociceptor sensitivity, but with different affinities. Such a mechanism could, for example, provide a gradual restoration of normal sensitivities among the population of nociceptors during wound healing, as the concentration of such modulators gradually returned to baseline. Such a graded response might be coupled to, or even determine the activation thresholds of different subsets of nociceptors. Another novel gene family isolated in this screen, *drg-12* encodes a protein with two putative transmembrane segments. *Drg12* was identified from both mice (SEQ ID NO: 14) and in humans (SEQ ID NO: 29). *In situ* hybridization indicates that, like the *mrg* genes, this gene is also specifically expressed in a subset of DRG sensory neurons. As it is a membrane protein it may also be involved in signaling by these neurons. Although there are no obvious homologies between this protein and other known proteins, it is noteworthy that two purinergic receptors specifically expressed in nociceptive sensory neurons ( $P_2X_2$  and  $P_2X_3$ ) have a similar bipartite transmembrane topology. Therefore it is likely that the family *drg-12* also encodes a receptor or ion channel involved in nociceptive sensory transduction or its modulation.

25           The proteins of the invention can serve as therapeutics and as a target for agents that modulate their expression or activity, for example in the treatment of chronic intractable pain and neuropathic pain. For example, agents may be identified which modulate biological processes associated with nociception such as the reception, transduction and transmission of pain signals.

## II. Specific Embodiments

### A. Definitions

Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. See, e.g. Singleton *et al.*, *Dictionary of Microbiology and Molecular Biology 2nd ed.*, J. Wiley & Sons (New York, NY 1994); Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, Cold Springs Harbor Press (Cold Springs Harbor, NY 1989). For purposes of the present invention, the following terms are defined below.

As used herein, the “protein” or “polypeptide” refers, in part, to a protein that has the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109. The terms also refer to naturally occurring allelic variants and proteins that have a slightly different amino acid sequence than those specifically recited above. Allelic variants, though possessing a slightly different amino acid sequence than those recited above, will still have the same or similar biological functions associated with the protein.

Identity or homology with respect to amino acid sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity (see section B for the relevant parameters). Fusion proteins, or N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

Proteins can be aligned using CLUSTALW (Thompson *et al.* Nucleic Acids Res 22:4673-80 (1994)) and homology or identity at the nucleotide or amino acid sequence level may be determined by BLAST (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs blastp, blastn, blastx, tblastn and tblastx (Karlin, *et al.* *Proc. Natl. Acad. Sci. USA* 87: 2264-2268 (1990) and Altschul, S. F. *J. Mol. Evol.* 36: 290-300 (1993), fully incorporated by reference) which are tailored for



sequence similarity searching. The approach used by the BLAST program is to first consider similar segments between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.* (Nature Genetics 6: 119-129 (1994)) which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (*i.e.*, the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the BLOSUM62 matrix (Henikoff, *et al.* Proc. Natl. Acad. Sci. USA 89: 10915-10919 (1992), fully incorporated by reference). For blastn, the scoring matrix is set by the ratios of M (*i.e.*, the reward score for a pair of matching residues) to N (*i.e.*, the penalty score for mismatching residues), wherein the default values for M and N are 5 and -4, respectively. Four blastn parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent Blastp parameter settings were Q=9; R=2; wink=1; and gapw=32. A Bestfit comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

“Variants” are biologically active polypeptides having an amino acid sequence which differs from the sequence of a native sequence polypeptide of the present invention, such as that shown in FIG. 1 for mrg3 (SEQ ID NO: 2), by virtue of an insertion, deletion, modification and/or substitution of one or more amino acid residues within the native sequence. Variants include peptide fragments of at least 5 amino acids, preferably at least 10 amino acids, more preferably at least 15 amino acids, even more preferably at least 20 amino acids that retain a biological activity of the corresponding native sequence polypeptide. Variants also include polypeptides wherein one or more amino acid residues are added at the N- or C-terminus of, or within, a native sequence. Further, variants also include polypeptides where a number of amino

acid residues are deleted and optionally substituted by one or more different amino acid residues.

As used herein, a “conservative variant” refers to alterations in the amino acid sequence that do not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

As used herein, the “family of proteins” related to the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109 includes proteins that have been isolated from the dorsal root ganglia of organisms in addition to mice and humans. The methods used to identify and isolate other members of the family of proteins related to these proteins, such as the disclosed mouse and human proteins, are described below.

Unless indicated otherwise, the term “Mrg” when used herein refers to any one or more of the mammalian mas-related gene (Mrg) receptors (*i.e.* MrgA1-8, MrgB, MrgC, MrgD, MrgE, MrgX1-4 and any other members of the mas-related gene (Mrg) family now known or identified in the future), including native sequence mammalian, such as murine or human, Mrg receptors, Mrg receptor variants; Mrg receptor extracellular domain; and chimeric Mrg receptors (each of which is defined herein). The term specifically includes native sequence murine Mrg receptors of the MrgA family, such as SEQ ID NOs: 2, 4, 6, 12, 21, 23, 25, 27, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75; native sequence murine Mrg receptors of the MrgB family, such as SEQ ID NOs: 39, 41, 43, 45, 47, 77, 79, 81, 83, 85, 87, and 89; native sequence murine Mrg receptors of the MrgC family, such as SEQ ID NOs: 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109; native sequence murine Mrg receptors of the MrgD family, such as SEQ ID NO: 49; native sequence murine Mrg receptors of the MrgE family, such as SEQ ID

NO: 51; their human homologues, and the native sequence human Mrg receptors termed "MrgX" of SEQ ID NOs: 16, 18, 31 and 33.

The terms "mas-related gene", "mrg" and "Mrg" are used interchangeably herein. Further, the terms mrg3, MrgA1 and mMrgA1 are used interchangeably, as are the terms mrg4, MrgA2 and mMrgA2, the terms mrg5, MrgA3 and mMrgA3, the terms mrg8, MrgA4 and mMrgA4, the terms mrg9, MrgA5 and mMrgA5, the terms mrg10, MrgA6 and mMrgA6, the terms mrg11, MrgA7 and mMrgA7, the terms mrg12, MrgA8 and mMrgA8, the terms human1, MrgX1 and hMrgX1, the terms human2, MrgX2 and hMrgX2, the terms human 4, MrgX3 and hMrgX3, and the terms human5, MrgX4 and hMrgX4. These terms all refer to native sequence Mrg proteins as described herein as well as functional derivatives, including amino acid sequence variants thereof.

A "native" or "native sequence" Mrg or drg-12 receptor has the amino acid sequence of a naturally occurring Mrg or drg-12 receptor in any mammalian species (including humans), irrespective of its mode of preparation. Accordingly, a native or native sequence Mrg or drg-12 receptor may be isolated from nature, produced by techniques of recombinant DNA technology, chemically synthesized, or produced by any combinations of these or similar methods. Native Mrg and drg-12 receptors specifically include polypeptides having the amino acid sequence of naturally occurring allelic variants, isoforms or spliced variants of these receptors, known in the art or hereinafter discovered.

The "extracellular domain" (ECD) is a form of the Mrg or drg-12 receptor which is essentially free of the transmembrane and cytoplasmic domains, i.e., has less than 1% of such domains, preferably 0.5 to 0% of such domains, and more preferably 0.1 to 0% of such domains. Ordinarily, the ECD will have an amino acid sequence having at least about 60% amino acid sequence identity with the amino acid sequence of one or more of the ECDs of a native Mrg or drg-12 protein, for example as indicated in FIG. 1A for mrg3 (E1, E2 etc...), preferably at least about 65%, more preferably at least about 75%, even more preferably at least about 80%, even more preferably at least about 90%, with increasing preference of 95%, to at least 99% amino acid sequence identity, and finally to 100% identity, and thus includes polypeptide variants as defined below.

5 The first predicted extracellular domain (ECD1) comprises approximately amino acids 1 to 21 for MrgA1, 1 to 21 for MrgA2, 1 to 21 for MrgA3, 1 to 21 for MrgA4, 1 to 3 for MrgA5, 1 to 17 for MrgA6, 1 to 21 for MrgA7 and 1 to 21 for MrgA8. The second predicted extracellular domain (ECD2) comprises approximately amino acids 70 to 87 for MrgA1, 70 to 88 for MrgA2, 70 to 88 for MrgA3, 70 to 88 for MrgA4, 52 to 70 for MrgA5, 66 to 84 for MrgA6, 70 to 88 for MrgA7 and 70 to 88 for MrgA8. The third predicted extracellular domain (ECD3) comprises approximately amino acids 149 to 160 for MrgA1, 150 to 161 for MrgA2, 150 to 161 for MrgA3, 150 to 161 for MrgA4, 132 to 144 for MrgA5, 146 to 157 for MrgA6, 150 to 161 for MrgA7 and 150 to 161 for MrgA8. The fourth predicted extracellular domain (ECD4) comprises approximately amino acids 222 to 2244 for MrgA1, 223 to 245 for MrgA2, 223 to 242 for MrgA3, 223 to 245 for MrgA4, 205 to 225 for MrgA5, 219 to 241 for MrgA6, 223 to 245 for MrgA7 and 223 to 245 for MrgA8.

15 The term “*drg-12*” when used herein refers to any one or more of the mammalian *drg-12* receptors now known or identified in the future, including native sequence mammalian, such as murine or human, *drg-12* receptors, *drg-12* receptor variants; *drg-12* receptor extracellular domain; and chimeric *drg-12* receptors (each of which is defined herein). The term specifically includes native sequence murine *drg-12* receptor, such as SEQ ID NO: 14, and any human homologues, such as human *drg-12* (SEQ ID NO: 29).

20 As used herein, “nucleic acid” is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to such a nucleic acid and remains stably bound to it under appropriate stringency conditions, exhibits at least about 50%, 60%, 70%, 75%, 85%, 25 90% or 95% nucleotide sequence identity across the open reading frame, or encodes a polypeptide sharing at least about 50%, 60%, 70% or 75% sequence identity, preferably at least about 80%, and more preferably at least about 85%, and even more preferably at least about 90 or 95% or more identity with the peptide sequences. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as 30 nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary

nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

5           As used herein, the terms nucleic acid, polynucleotide and nucleotide are interchangeable and refer to any nucleic acid, whether composed of phosphodiester linkages or modified linkages such as phosphotriester, phosphoramidate, siloxane, carbonate, carboxymethylester, acetamidate, carbamate, thioether, bridged phosphoramidate, bridged methylene phosphonate, bridged phosphoramidate, bridged phosphoramidate, bridged methylene phosphonate, phosphorothioate, methylphosphonate, phosphorodithioate, bridged phosphorothioate or sultone linkages, and combinations of such linkages.

15           The terms nucleic acid, polynucleotide and nucleotide also specifically include nucleic acids composed of bases other than the five biologically occurring bases (adenine, guanine, thymine, cytosine and uracil). For example, a polynucleotide of the invention might contain at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyl-uracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 20           1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5N-methoxycarboxymethyluracil, 5-methoxyuracil, 25           2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

Furthermore, a polynucleotide used in the invention may comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C., or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is use of 50% formamide, 5 x SSC (0.75M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal.

As used herein, a nucleic acid molecule is said to be "isolated" when the nucleic acid molecule is substantially separated from contaminant nucleic acid molecules encoding other polypeptides.

As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. For instance, fragments which encode peptides corresponding to predicted antigenic regions may be prepared (see Figures 3 and 4). If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming (see the discussion in Section H).

Highly related gene homologs are polynucleotides encoding proteins that have at least about 60% amino acid sequence identity with the amino acid sequence of a naturally occurring native sequence polynucleotide of the invention, such as MrgA1 (SEQ ID NO: 2), preferably at least about 65%, 70%, 75%, 80%, with increasing

preference of at least about 85% to at least about 99% amino acid sequence identity, in 1% increments.

The term "mammal" is defined as an individual belonging to the class Mammalia and includes, without limitation, humans, domestic and farm animals, and zoo, sports, or pet animals, such as sheep, dogs, horses, cats or cows. Preferably, the mammal herein is human.

"Functional derivatives" include amino acid sequence variants, and covalent derivatives of the native polypeptides as long as they retain a qualitative biological activity of the corresponding native polypeptide.

By "Mrg ligand" is meant a molecule which specifically binds to and preferably activates an Mrg receptor. Examples of Mrg ligands include, but are not limited to RF-amide neuropeptides, such as FMRF, FLRF, NPAF, NPFF, and RFRP-1 for MrgA receptors, such as MrgA1. The ability of a molecule to bind to Mrg can be determined, for example, by the ability of the putative ligand to bind to membrane fractions prepared from cells expressing Mrg.

Similarly, a *drg-12* ligand is a molecule which specifically binds to and preferably activates a *drg-12* receptor.

A "chimeric" molecule is a polypeptide comprising a full-length polypeptide of the present invention, a variant, or one or more domains of a polypeptide of the present invention fused or bonded to a heterologous polypeptide. The chimeric molecule will generally share at least one biological property in common with a naturally occurring native sequence polypeptide. An example of a chimeric molecule is one that is epitope tagged for purification purposes. Another chimeric molecule is an immunoadhesin.

The term "epitope-tagged" when used herein refers to a chimeric polypeptide comprising Mrg or drg-12 fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with the biological activity of the Mrg or drg-12. The tag polypeptide preferably is fairly unique so that the antibody against it does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8-50 amino acid residues (preferably between about 9-30 residues). Preferred are poly-histidine sequences, which

bind nickel, allowing isolation of the tagged protein by Ni-NTA chromatography as described (See, e.g., Lindsay *et al. Neuron* 17:571-574 (1996)).

"Agonists" are molecules or compounds that stimulate one or more of the biological properties of a polypeptide of the present invention. These may include, but are not limited to, small organic and inorganic molecules, peptides, peptide mimetics and agonist antibodies.

The term "antagonist" is used in the broadest sense and refers to any molecule or compound that blocks, inhibits or neutralizes, either partially or fully, a biological activity mediated by a receptor of the present invention by preventing the binding of an agonist. Antagonists may include, but are not limited to, small organic and inorganic molecules, peptides, peptide mimetics and neutralizing antibodies.

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein. In some instances, isolated proteins of the invention will have been separated or purified from many cellular constituents, but will still be associated with cellular membrane fragments or membrane constituents.

Thus, "isolated Mrg" and "isolated drg-12" means Mrg or drg-12 polypeptide, respectively, that has been purified from a protein source or has been prepared by recombinant or synthetic methods and purified. Purified Mrg or drg-12 is substantially free of other polypeptides or peptides. "Substantially free" here means less than about 5%, preferably less than about 2%, more preferably less than about 1%, even more preferably less than about 0.5%, most preferably less than about 0.1% contamination with other source proteins.

"Essentially pure" protein means a composition comprising at least about 90% by weight of the protein, based on total weight of the composition, preferably at least about 95% by weight, more preferably at least about 90% by weight, even more preferably at least about 95% by weight. "Essentially homogeneous" protein means a composition comprising at least about 99% by weight of protein, based on total weight of the composition.



"Biological property" is a biological or immunological activity, where biological activity refer to a biological function (either inhibitory or stimulatory) caused by a native sequence or variant polypeptide molecule herein, other than the ability to induce the production of an antibody against an epitope within such polypeptide, where the latter property is referred to as immunological activity. Biological properties specifically include the ability to bind a naturally occurring ligand of the receptor molecules herein, preferably specific binding, and even more preferably specific binding with high affinity.

"Antibodies" (Abs) and "immunoglobulins" (Igs) are glycoproteins having the same structural characteristics. While antibodies exhibit binding specificity to a specific antigen, immunoglobulins include both antibodies and other antibody-like molecules that lack antigen specificity. Polypeptides of the latter kind are, for example, produced at low levels by the lymph system and at increased levels by myelomas.

"Native antibodies" and "native immunoglobulins" are usually heterotetrameric glycoproteins, composed of two identical light (L) chains and two identical heavy (H) chains. Each light chain is linked to a heavy chain by one covalent disulfide bond. while The number of disulfide linkages varies among the heavy chains of different immunoglobulin isotypes. Each heavy and light chain also has regularly spaced intra-chain disulfide bridges. Each heavy chain has at one end a variable domain ( $V_H$ ) followed by a number of constant domains. Each light chain has a variable domain at one end ( $V_L$ ) and a constant domain at its other end. The constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light- chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light- and heavy-chain variable domains.

The term "antibody" herein is used in the broadest sense and specifically covers human, non-human (e.g. murine) and humanized monoclonal antibodies (including full length monoclonal antibodies), polyclonal antibodies, multi-specific antibodies (e.g., bispecific antibodies), and antibody fragments so long as they exhibit the desired biological activity.

"Antibody fragments" comprise a portion of a full-length antibody, generally the antigen binding or variable domain thereof. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies; single-chain antibody molecules; and multi-specific antibodies formed from antibody fragments.

5           The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of antibodies wherein the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific and are directed against a single antigenic site. In addition, monoclonal antibodies may be made by any  
10           method known in the art. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler *et al.*, *Nature* 256:495 (1975), or may be made by recombinant DNA methods (see, *e.g.*, U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in  
15           Clackson *et al.*, *Nature* 352:624-628 (1991) and Marks *et al.*, *J. Mol. Biol.* 222:581-597 (1991), for example.

          The monoclonal antibodies herein specifically include "chimeric" antibodies (immunoglobulins) in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular  
20           species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass. Fragments of chimeric antibodies are also included provided they exhibit the desired biological activity (U.S. Patent No. 4,816,567; and Morrison *et al.*, *Proc. Natl. Acad. Sci. USA* 81:6851-6855 (1984)).  
25           

          "Humanized" forms of non-human (*e.g.*, murine) antibodies are antibodies that contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies are generally human immunoglobulins in which hypervariable region residues are replaced by hypervariable region residues from a non-human species such  
30           as mouse, rat, rabbit or non-human primate having the desired specificity, affinity, and capacity. Framework region (FR) residues of the human immunoglobulin may be

replaced by corresponding non-human residues. In addition, humanized antibodies may comprise residues that are not found in either the recipient antibody or in the donor antibody. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable regions correspond to those of a non-human immunoglobulin and all or substantially all of the FRs are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones *et al.*, *Nature* 321:522-525 (1986); Reichmann *et al.*, *Nature* 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992).

The term "epitope" is used to refer to binding sites for (monoclonal or polyclonal) antibodies on protein antigens.

By "agonist antibody" is meant an antibody which is a ligand for a receptor of the invention and thus, able to activate and/or stimulate one or more of the effector functions of native sequence Mrg or drg-12.

By "neutralizing antibody" is meant an antibody molecule as herein defined which is able to block or significantly reduce an effector function of a polypeptide of the invention. For example, a neutralizing antibody may inhibit or reduce Mrg or drg-12 activation by a known ligand.

The term "Mrg immunoadhesin" refers to a chimeric molecule that comprises at least a portion of an Mrg or drg-12 molecule (native or variant) and an immunoglobulin sequence. The immunoglobulin sequence preferably, but not necessarily, is an immunoglobulin constant domain. Immunoadhesins can possess many of the properties of human antibodies. Since immunoadhesins can be constructed from a human protein sequence with a desired specificity linked to an appropriate human immunoglobulin hinge and constant domain (Fc) sequence, the binding specificity of interest can be achieved using entirely human components. Such immunoadhesins are minimally immunogenic to the patient, and are safe for chronic or repeated use. If the two arms of the immunoadhesin structure have different specificities, the immunoadhesin is called a "bispecific immunoadhesin" by analogy to bispecific antibodies.

As used herein, "treatment" is a clinical intervention made in response to a disease, disorder or physiological condition manifested by a patient. The aim of treatment includes the alleviation or prevention of symptoms, slowing or stopping the progression or worsening of a disease, disorder, or condition and the remission of the disease, disorder or condition. "Treatment" refers to both therapeutic treatment and prophylactic or preventative measures. Those in need of treatment include those already affected by a disease or disorder or undesired physiological condition as well as those in which the disease or disorder or undesired physiological condition is to be prevented. Specifically, treatment may alleviate pain, including pain resulting from an existing condition or disorder, or to prevent pain in situations where pain is likely to be experienced.

In the methods of the present invention, the term "control" and grammatical variants thereof, are used to refer to the prevention, partial or complete inhibition, reduction, delay or slowing down of an unwanted event, such as the presence or onset of pain.

The term "effective amount" refers to an amount sufficient to effect beneficial or desirable clinical results. An effective amount of an agonist or antagonist is an amount that is effective to treat a disease, disorder or unwanted physiological condition.

"Pain" is a sensory experience perceived by nerve tissue distinct from sensations of touch, pressure, heat and cold. The range of pain sensations, as well as the variation of perception of pain by individuals, renders a precise definition of pain near impossible. In the context of the present invention, "pain" is used in the broadest possible sense and includes nociceptive pain, such as pain related to tissue damage and inflammation, pain related to noxious stimuli, acute pain, chronic pain, and neuropathic pain.

"Acute pain" is often short-lived with a specific cause and purpose; generally produces no persistent psychological reactions. Acute pain can occur during soft tissue injury, and with infection and inflammation. It can be modulated and removed by treating its cause and through combined strategies using analgesics to treat the pain and antibiotics to treat the infection.

“Chronic pain” is distinctly different from and more complex than acute pain. Chronic pain has no time limit, often has no apparent cause and serves no apparent biological purpose. Chronic pain can trigger multiple psychological problems that confound both patient and health care provider, leading to feelings of helplessness and hopelessness. The most common causes of chronic pain include low-back pain, headache, recurrent facial pain, pain associated with cancer and arthritis pain.

The pain is termed “neuropathic” when it is taken to represent neurologic dysfunction. “Neuropathic pain” has a complex and variable etiology. It is typically characterized by hyperalgesia (lowered pain threshold and enhanced pain perception) and by allodynia (pain from innocuous mechanical or thermal stimuli). Neuropathic pain is usually chronic and tends not to respond to the same drugs as “normal pain” (nociceptive pain), therefore, its treatment is much more difficult. Neuropathic pain may develop whenever nerves are damaged, by trauma, by disease such as diabetes, herpes zoster, or late-stage cancer, or by chemical injury (e.g., as an untoward consequence of agents including the false-nucleotide anti-HIV drugs). It may also develop after amputation (including mastectomy). Examples of neuropathic pain include monoradiculopathies, trigeminal neuralgia, postherpetic neuralgia, complex regional pain syndromes and the various peripheral neuropathies. This is in contrast with “normal pain” or “nociceptive pain,” which includes normal post-operative pain, pain associated with trauma, and chronic pain of arthritis.

“Peripheral neuropathy” is a neurodegenerative disorder that affects the peripheral nerves, most often manifested as one or a combination of motor, sensory, sensorimotor, or autonomic dysfunction. Peripheral neuropathies may, for example, be characterized by the degeneration of peripheral sensory neurons, which may result from a disease or disorder such as diabetes (diabetic neuropathy), alcoholism and acquired immunodeficiency syndrome (AIDS), from therapy such as cytostatic drug therapy in cancer, or from genetic predisposition. Genetically acquired peripheral neuropathies include, for example, Krabbe’s disease, Metachromatic leukodystrophy, and Charcot-Marie-Tooth (CMT) Disease. Peripheral neuropathies are often accompanied by pain.

“Pharmaceutically acceptable” carriers, excipients, or stabilizers are ones which are nontoxic to the cell or mammal being exposed thereto at the dosages and

concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution such as phosphate buffer or citrate buffer. The physiologically acceptable carrier may also comprise one or more of the following: antioxidants including ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, such as serum albumin, gelatin, immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids, carbohydrates including glucose, mannose, or dextrans, chelating agents such as EDTA, sugar alcohols such as mannitol or sorbitol, salt-forming counterions such as sodium, and nonionic surfactants such as Tween™, polyethylene glycol (PEG), and Pluronics™.

"Peptide mimetics" are molecules which serve as substitutes for peptides in interactions with the receptors of the present invention (Morgan *et al.*, *Ann. Reports Med. Chem.* 24:243-252 (1989)). Peptide mimetics, as used herein, include synthetic structures that retain the structural and functional features of a peptide. Peptide mimetics may or may not contain amino acids and/or peptide bonds. The term, "peptide mimetics" also includes peptoids and oligopeptoids, which are peptides or oligomers of N-substituted amino acids (Simon *et al.*, *Proc. Natl. Acad. Sci. USA* 89:9367-9371 (1972)). Further included as peptide mimetics are peptide libraries, which are collections of peptides designed to be of a given amino acid length and representing all conceivable sequences of amino acids corresponding thereto.

#### **A. Proteins Expressed in Primary Sensory Neurons of Dorsal Root Ganglia**

The present invention provides isolated *mrg* and *drg-12* proteins, allelic variants of the proteins, and conservative amino acid substitutions of the proteins. Polypeptide sequences of several Mrg proteins of the present invention are provided in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109. Polypeptide sequences of several *drg-12* proteins of the present invention are provided in SEQ ID NOs: 14, 19 and 29.

The proteins of the present invention further include insertion, deletion or conservative amino acid substitution variants of the sequences set forth in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49,

51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109.

Ordinarily, the variants, allelic variants, the conservative substitution variants, and the members of the protein family, including corresponding homologues in other species, will have an amino acid sequence having at least about 50%, or about 60% to 75% amino acid sequence identity with the sequences set forth in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 or 109, more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% sequence identity with said sequences.

The proteins of the present invention include molecules having the amino acid sequence disclosed in SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 and 109; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues of the protein; amino acid sequence variants wherein one or more amino acid residues has been inserted N- or C-terminal to, or within, the disclosed coding sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by another residue. Such fragments, also referred to as peptides or polypeptides, may contain antigenic regions, functional regions of the protein identified as regions of the amino acid sequence which correspond to known protein domains, as well as regions of pronounced hydrophilicity. The regions are all easily identifiable by using commonly available protein sequence analysis software such as MACVECTOR™ (Oxford Molecular).

Contemplated variants further include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, rat, porcine, bovine, ovine, equine, human and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other

appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

Protein domains such as a ligand binding domain, an extracellular domain, a transmembrane domain (e.g. comprising seven membrane spanning segments and cytosolic loops or two membrane spanning domains and cytosolic loops), the transmembrane domain and a cytoplasmic domain and an active site may all be found in the proteins or polypeptides of the invention. Such domains are useful for making chimeric proteins and for *in vitro* assays of the invention.

Variations in native sequence proteins of the present invention or in various domains identified therein, can be made, for example, using any techniques known in the art. Variation can be achieved, for example, by substitution of at least one amino acid with any other amino acid in one or more of the domains of the protein. A change in the amino acid sequence of a protein of the invention as compared with a native sequence protein may be produced by a substitution, deletion or insertion of one or more codons encoding the protein. A comparison of the sequence of the Mrg or drg-12 polypeptide to be changed with that of homologous known protein molecules may provide guidance as to which amino acid residues may be inserted, substituted or deleted without affecting a desired biological activity. In particular, it may be beneficial to minimize the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

Polypeptide fragments are also useful in the methods of the present invention. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full-length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the Mrg or drg-12 polypeptide.



Mrg or *drg-12* fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized or generated by enzymatic digestion, such as by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues. Alternatively, the DNA encoding the protein may be digested with suitable restriction enzymes and the desired fragment isolated. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, Mrg or *drg-12* polypeptide fragments share at least one biological and/or immunological activity with a native Mrg or *drg-12* polypeptide, respectively.

In making amino acid sequence variants that retain the required biological properties of the corresponding native sequences, the hydropathic index of amino acids may be considered. For example, it is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score without significant change in biological activity. Thus, isoleucine, which has a hydropathic index of + 4.5, can generally be substituted for valine (+ 4.2) or leucine (+ 3.8), without significant impact on the biological activity of the polypeptide in which the substitution is made. Similarly, usually lysine (-3.9) can be substituted for arginine (-4.5), without the expectation of any significant change in the biological properties of the underlying polypeptide. Other considerations for choosing amino acid substitutions include the similarity of the side-chain substituents, for example, size, electrophilic character, charge in various amino acids. In general, alanine, glycine and serine; arginine and lysine; glutamate and aspartate; serine and threonine; and valine, leucine and isoleucine are interchangeable, without the expectation of any significant change in biological properties. Such substitutions are generally referred to as conservative amino acid substitutions, and are the preferred type of substitutions within the polypeptides of the present invention.

Non-conservative substitutions will entail exchanging a member of one class of amino acids for another class. Such substituted residues also may be introduced into

the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as site-directed mutagenesis, alanine scanning mutagenesis, and PCR mutagenesis. Site-directed mutagenesis (Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)), cassette mutagenesis (Wells et al., Gene, 34:315 (1985)), restriction selection mutagenesis (Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)) or other known techniques can be performed on cloned DNA to produce the Mrg or *drg-12* variant DNA.

Scanning amino acid analysis can be employed to identify one or more amino acids that can be replaced without a significant impact on biological activity. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is preferred because, in addition to being the most common amino acid, it eliminates the side-chain beyond the beta-carbon and is therefore less likely to alter the main-chain conformation of the variant (Cunningham and Wells, Science, 244: 1081-1085 (1989)). Further, alanine is frequently found in both buried and exposed positions (Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)). If alanine substitution does not yield adequate amounts of variation, an isoteric amino acid can be used.

As described below, members of the family of proteins can be used: 1) to identify agents which modulate at least one activity of the protein; 2) to identify binding partners for the protein, 3) as an antigen to raise polyclonal or monoclonal antibodies, 4) as a therapeutic target, 5) as diagnostic markers to specific populations of pain sensing neurons and 6) as targets for structure based ligand identification.

## **B. Nucleic Acid Molecules**

The present invention further provides nucleic acid molecules that encode the *mrg* or *drg-12* proteins having SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 or 109 and the related polypeptides herein described, preferably in isolated form. cDNAs encoding

eight full-length variants of Mrg receptors (mMrgA1-8) are provided in Figure 6A (SEQ ID NO: 1, 3, 5, 11, 20, 22, 24, 26).

Preferred molecules are those that hybridize under the above defined stringent conditions to the complement of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 20, 22, 24, 26 or 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 7274, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106 or 108 and which encode a functional peptide. Preferred hybridizing molecules are those that hybridize under the above conditions to the complement strand of the open reading frame or coding sequences of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 20, 22, 24, 26 or 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 7274, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106 or 108.

It is not intended that the methods of the present invention be limited by the source of the polynucleotide. The polynucleotide can be from a human or non-human mammal, derived from any recombinant source, synthesized *in vitro* or by chemical synthesis. The nucleotide may be DNA or RNA and may exist in a double-stranded, single-stranded or partially double-stranded form.

Nucleic acids useful in the present invention include, by way of example and not limitation, oligonucleotides such as antisense DNAs and/or RNAs; ribozymes; DNA for gene therapy; DNA and/or RNA chimeras; various structural forms of DNA including single-stranded DNA, double-stranded DNA, supercoiled DNA and/or triple-helix DNA; Z-DNA; and the like. The nucleic acids may be prepared by any conventional means typically used to prepare nucleic acids in large quantity. For example, DNAs and RNAs may be chemically synthesized using commercially available reagents and synthesizers by methods that are well-known in the art (*see, e.g.*, Gait, 1985, Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford, England).

Any mRNA transcript encoded by Mrg or drg-12 nucleic acid sequences may be used in the methods of the present invention, including in particular, mRNA transcripts resulting from alternative splicing or processing of mRNA precursors.

Nucleic acids having modified nucleoside linkages may also be used in the methods of the present invention. Modified nucleic acids may, for example, have

greater resistance to degradation. Such nucleic acids may be synthesized using reagents and methods that are well known in the art. For example, methods for synthesizing nucleic acids containing phosphonate phosphorothioate, phosphorodithioate, phosphoramidate methoxyethyl phosphoramidate, formacetal, thioformacetal, diisopropylsilyl, acetamidate, carbamate, dimethylene-sulfide (-CH<sub>2</sub>-S-CH<sub>2</sub>), dimethylene-sulfoxide (-CH<sub>2</sub>-SO-CH<sub>2</sub>), dimethylene-sulfone (-CH<sub>2</sub>-SO<sub>2</sub>-CH<sub>2</sub>), 2'-O-alkyl, and 2'-deoxy-2'-fluoro phosphorothioate internucleoside linkages are well known in the art.

In some embodiments of the present invention, the nucleotide used is an  $\alpha$ -anomeric nucleotide. An  $\alpha$ -anomeric nucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier *et al.*, 1987, Nucl. Acids Res. 15:6625-6641). The nucleotide may be a 2'-O-methylribonucleotide (Inoue *et al.*, 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987, FEBS Lett. 215:327-330).

Means for purifying the nucleic acids of the present invention are well known in the art and the skilled artisan will be able to choose the most appropriate method of purification for the particular circumstances. Such a choice may be made, in part, based on the size of the DNA, the amount to be purified and the desired purity. For example, the nucleic acids can be purified by reverse phase or ion exchange HPLC, size exclusion chromatography or gel electrophoresis.

Isolated or purified polynucleotides having at least 10 nucleotides (*i.e.*, a hybridizable portion) of an Mrg or drg-12 coding sequence or its complement may also be used in the methods of the present invention. In other embodiments, the polynucleotides contain at least 25 (continuous) nucleotides, 50 nucleotides, 100 nucleotides, 150 nucleotides, or 200 nucleotides of an Mrg coding sequence, or a full-length Mrg coding sequence. Nucleic acids can be single or double stranded. Additionally, the invention relates to polynucleotides that selectively hybridize to a complement of the foregoing coding sequences. In preferred embodiments, the polynucleotides contain at least 10, 25, 50, 100, 150 or 200 nucleotides or the entire length of an Mrg coding sequence.

Nucleotide sequences that encode a mutant of an Mrg protein, peptide fragments of Mrg, truncated forms of Mrg, and Mrg fusion proteins may also be useful in the methods of the present invention. Nucleotides encoding fusion proteins may include, but are not limited to, full length Mrg sequences, truncated forms of Mrg, or  
5 nucleotides encoding peptide fragments of Mrg fused to an unrelated protein or peptide, such as for example, a domain fused to an Ig Fc domain or fused to an enzyme such as a fluorescent protein or a luminescent protein which can be used as a marker.

Furthermore, polynucleotide variants that have been generated, at least in part, by some form of directed evolution, such as gene shuffling or recursive sequence  
10 recombination may be used in the methods of the present invention. For example, using such techniques novel sequences can be generated encoding proteins similar to Mrg or drg-12 but having altered functional or structural characteristics.

Highly related gene homologs of the Mrg encoding polynucleotide sequences described above may also be useful in the present invention. Highly related homologs  
15 can encode proteins sharing functional activities with Mrg proteins.

The present invention further provides fragments of the encoding nucleic acid molecule. Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of  
20 the invention, can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, (J. Am. Chem. Soc. 103:3185-3191, 1981) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides  
25 to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to,  
30 biotin, radiolabeled nucleotides and the like. A skilled artisan can readily employ any such label to obtain labeled variants of the nucleic acid molecules of the invention.

Any nucleotide sequence which encodes the amino acid sequence of a protein of the invention can be used to generate recombinant molecules which direct the expression of the protein, as described in more detail below. In addition, the methods of the present invention may also utilize a fusion polynucleotide comprising an Mrg or drg-12 coding sequence and a second coding sequence for a heterologous protein.

### C. Isolation of Other Related Nucleic Acid Molecules

As described above, the identification and characterization of a nucleic acid molecule encoding an *mrg* or *drg-12* protein allows a skilled artisan to isolate nucleic acid molecules that encode other members of the same protein family in addition to the sequences herein described

Essentially, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 or 109 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as a lambda gt11 library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the protein.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the Mrg protein family from cells derived from any mammalian organism, particularly cells believed to express Mrg proteins. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives. Oligonucleotides

corresponding to either the 5' or 3' terminus of the coding sequence may be used to obtain longer nucleotide sequences.

It may be necessary to screen multiple cDNA libraries to obtain a full-length cDNA. In addition, it may be necessary to use a technique such as the RACE (Rapid Amplification of cDNA Ends) technique to obtain the complete 5' terminal coding region. RACE is a PCR-based strategy for amplifying the 5' end of incomplete cDNAs. To obtain the 5' end of the cDNA, PCR is carried out on 5'-RACE-Ready cDNA using an anchor primer and a 3' primer. A second PCR is then carried out using the anchored primer and a nested 3' primer. Once a full length cDNA sequence is obtained, it may be translated into amino acid sequence and examined for identifiable regions such as a continuous open reading frame flanked by translation initiation and termination sites, a potential signal sequence and finally overall structural similarity to the protein sequences disclosed herein.

Related nucleic acid molecules may also be retrieved by using pairs of oligonucleotide primers in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. The oligonucleotide primers may be degenerate oligonucleotide primer pools designed on the basis of the protein coding sequences disclosed herein. The template for the reaction may be cDNA obtained by reverse transcription (RT) of mRNA prepared from, for example, human or non-human cell lines or tissues known or suspected to express an Mrg or drg-12 gene allele, such as DRG tissue. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of an Mrg or drg-12 coding sequence. The PCR fragment may then be used to isolate a full-length cDNA clone by a variety of methods. For example, the amplified fragment may be labeled and used to screen a cDNA library. Alternatively, the labeled fragment may be used to isolate genomic clones via the screening of a genomic library.

PCR technology may also be utilized to isolate full-length cDNA sequences. RNA may be isolated, from an appropriate cellular or tissue source, such as dorsal root

ganglion (DRG) and an RT reaction may be carried out using an oligonucleotide primer specific for the most 5' end of the amplified fragment to prime first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines in a terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. This allows isolation of cDNA sequences upstream of the amplified fragment.

Nucleic acid molecules encoding other members of the *mrg* and *drg-12* families may also be identified in existing genomic or other sequence information using any available computational method, including but not limited to: PSI-BLAST (Altschul, *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402); PHI-BLAST (Zhang, *et al.* (1998), *Nucleic Acids Res.* 26:3986-3990), 3D-PSSM (Kelly *et al.* *J. Mol. Biol.* 299(2): 499-520 (2000)); and other computational analysis methods (Shi *et al.* *Biochem. Biophys. Res. Commun.* 262(1):132-8 (1999) and Matsunami *et. al.* *Nature* 404(6778):601-4 (2000).

A cDNA clone of a mutant or allelic variant of an Mrg or drg-12 gene may also be isolated. A possible source of a mutant or variant protein is tissue known to express Mrg or drg-12, such as DRG tissue, obtained from an individual putatively carrying a mutant or variant form of Mrg or drg-12. Such an individual may be identified, for example, by a demonstration of increased or decreased responsiveness to painful stimuli. In one embodiment, a mutant or variant Mrg or drg-12 gene may be identified by PCR. The first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from the tissue putatively carrying a variant and extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant Mrg allele to that of the normal Mrg allele, the mutation(s) responsible for any loss or alteration of function of the mutant Mrg gene product can be ascertained.

Alternatively, a genomic library can be constructed using DNA obtained from an individual suspected of or known to carry a mutant Mrg allele, or a cDNA library



can be constructed using RNA from a tissue known, or suspected, to express a mutant Mrg allele. An unimpaired Mrg gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant Mrg allele in such libraries. Clones containing the mutant Mrg gene sequences may then be purified and subjected to sequence analysis according to methods well known to those of skill in the art.

Additionally, an expression library can be constructed utilizing cDNA synthesized from, for example, RNA isolated from a tissue known, or suspected, to express a mutant Mrg allele in an individual suspected of carrying such a mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal Mrg gene product, as described, below.

#### **D. Recombinant DNA molecules containing a Nucleic Acid Molecule**

The present invention further provides recombinant DNA molecules (rDNAs) that contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in situ*. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd edition, 1989; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

Thus the present invention also contemplates DNA vectors that contain any of the Mrg or drg-12 coding sequences and/or their complements, optionally associated with a regulatory element that directs the expression of the coding sequences. The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Both cloning and expression vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. In cloning vectors this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 $\mu$  plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

In addition to being capable of replication in at least one class of organism most expression vectors can be transfected into another organism for expression. For example, a vector is replicated in *E. coli* and then the same vector is transfected into yeast or mammalian cells for expression.

DNA may also be amplified by insertion into the host genome. For example, transfection of *Bacillus* with a vector comprising a DNA sequence complementary to a *Bacillus* genomic sequence results in homologous recombination with the genome and insertion of the DNA from the vector. One disadvantage to this type of system is that the recovery of genomic DNA encoding the protein of interest is more complex than that of an exogenously replicated vector because restriction enzyme digestion is required to excise the DNA.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable

marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

5 Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences that are compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from BioRad Laboratories, (Richmond, CA), pPL and pKK223 available from Pharmacia (Piscataway, NJ).

15 Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form rDNA molecules that contain a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), eukaryotic viral vectors such as adenoviral or retroviral vectors, and the like eukaryotic expression vectors.

20 Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. This gene encodes a factor necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, complement auxotrophic deficiencies, or supply critical nutrients withheld from the media. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, *i.e.*, the neomycin phosphotransferase (neo) gene. (Southern *et al.*, J. Mol. Anal. Genet. 1:327-341, 1982.) The selectable

25  
30

marker can optionally be present on a separate plasmid and introduced by co-transfection.

In one example of a selection system, mammalian cell transformants are placed under selection pressure such that only the transformants are able to survive by virtue of having taken up the vector(s). Selection pressure is imposed by progressively increasing the concentration of selection agent in the culture medium, thereby stimulating amplification of both the selection gene and the DNA that encodes the desired protein. Amplification is the process by which genes in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Increased quantities of the desired protein, such as Mrg, are synthesized from the amplified DNA. Examples of amplifiable genes include DHFR, thymidine kinase, metallothionein-I and -II, adenosine deaminase, and ornithine decarboxylase.

Thus in one embodiment Chinese hamster ovary (CHO) cells deficient in DHFR activity are prepared and propagated as described by Urlaub *et al.*, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980). The CHO cells are then transformed with the DHFR selection gene and transformants are identified by culturing in a culture medium that contains methotrexate (Mtx), a competitive antagonist of DHFR. The transformed cells are then exposed to increased levels of methotrexate. This leads to the synthesis of multiple copies of the DHFR gene, and, concomitantly, multiple copies of other DNA comprising the expression vectors, such as the DNA encoding the protein of interest, for example DNA encoding Mrg.

Alternatively, host cells can be transformed or co-transformed with DNA sequences encoding a protein of interest such as Mrg, wild-type DHFR protein, and another selectable marker such as aminoglycoside 3'-phosphotransferase (APH). The transformants can then be selected by growth in medium containing a selection agent for the selectable marker such as an aminoglycosidic antibiotic, *e.g.*, kanamycin, neomycin, or G418.

As mentioned above, expression and cloning vectors usually contain a promoter that is recognized by the host organism and is operably linked to the nucleic acid encoding the protein of interest. Promoters are untranslated sequences located upstream

(5') to the start codon of a structural gene (generally within about 100 to 1000 bp) and control the transcription and translation of the particular nucleic acid sequence, such as an Mrg nucleic acid sequence, to which they are operably linked. Promoters may be inducible or constitutive. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as a change in temperature. Many different promoters are well known in the art, as are methods for operably linking the promoter to the DNA encoding the protein of interest. Both the native Mrg or drg-12 promoter sequence and many heterologous promoters may be used to direct amplification and/or expression of the Mrg or drg-12 DNA. However, heterologous promoters are preferred, as they generally permit greater transcription and higher yields of the desired protein as compared to the native promoter.

Promoters suitable for use with prokaryotic hosts include, for example, the  $\beta$ -lactamase and lactose promoter systems (Chang *et al.*, *Nature*, 275:615 (1978); Goeddel *et al.*, *Nature*, 281:544 (1979)). However, other bacterial promoters are well known in the art and are suitable. Promoters for use in bacterial systems also will contain a Shine-Delgarno (S.D.) sequence operably linked to the DNA encoding the protein of interest.

Promoter sequences that can be used in eukaryotic cells are also well known. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the transcription initiation site. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AATAAA sequence that may be the signal for addition of the poly-A tail to the 3' end of the coding sequence. All of these sequences may be inserted into eukaryotic expression vectors.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, *J. Biol. Chem.*, 255:2073 (1980)) or other glycolytic enzymes (Hess *et al.*, *J. Adv. Enzyme Reg.*, 7:149 (1968); Holland, *Biochemistry*, 17:4900 (1978)).

Inducible promoters for use with yeast are also well known and include the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657. Yeast enhancers also are advantageously used with yeast promoters.

*Mrg* or *drg-12* transcription from vectors in mammalian host cells may also be controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus, bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the native sequence, provided such promoters are compatible with the host cell systems.

Transcription may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about 10 to 300 bp in length, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin). Preferably an enhancer from a eukaryotic cell virus will be used. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the protein-encoding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. These sequences are often found in the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs and are well known in the art.

Plasmid vectors containing one or more of the components described above are readily constructed using standard techniques well known in the art.

For analysis to confirm correct sequences in plasmids constructed, the plasmid may be replicated in *E. coli*, purified, and analyzed by restriction endonuclease digestion, and/or sequenced by conventional methods.

Particularly useful in the preparation of proteins of the present invention are expression vectors that provide for transient expression in mammalian cells of DNA encoding Mrg or drg-12. Transient expression involves the use of an expression vector that is able to replicate efficiently in a host cell, such that the host cell accumulates many copies of the expression vector and, in turn, synthesizes high levels of a the polypeptide encoded by the expression vector. Sambrook *et al.*, *supra*, pp. 16.17 - 16.22. Transient expression systems allow for the convenient positive identification of polypeptides encoded by cloned DNAs, as well as for the screening of such polypeptides for desired biological or physiological properties. Thus, transient expression systems are particularly useful in the invention for purposes of identifying biologically active analogs and variants of the polypeptides of the invention and for identifying agonists and antagonists thereof.

Other methods, vectors, and host cells suitable for adaptation to the synthesis of Mrg or drg-12 in recombinant vertebrate cell culture are well known in the art and are readily adapted to the specific circumstances.

#### **E. Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule**

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic but is preferably eukaryotic.

Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and

mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells (NIH/3T3) available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), HEK293 cells and the like eukaryotic tissue culture cell lines.

Propagation of vertebrate cells in culture is a routine procedure. See, e.g., *Tissue Culture*, Academic Press, Kruse and Patterson, editors (1973). Additional examples of useful mammalian host cell lines that can be readily cultured are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51).

*Xenopus* oocytes may also be directly injected with RNA capable of expressing either the *mrg* or *drg-12* proteins by standard procedures (see Tominaga *et al.* Jpn J. Pharmacol. 83(1):20-4 (2000); Tominaga *et al.* Neuron 21(3):531-43 (1998) and Bisogno *et al.* Biochem. Biophys. Res. Commun. 262(1):275-84 (1999)).

Examples of invertebrate cells that can be used as hosts include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells are known in the art and may be utilized in the methods of the present invention. In addition, plant cell cultures are known and may be transfected, for example, by incubation with *Agrobacterium tumefaciens*, which has been manipulated to contain Mrg or drg-12 encoding DNA.

Any prokaryotic host can be used to express a rDNA molecule encoding a protein or a protein fragment of the invention. Suitable prokaryotes include eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B.*



*licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. The preferred prokaryotic host is *E. coli*. In addition, it is preferably that the host cell secrete minimal amounts of proteolytic enzymes.

5 In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for Mrg- or drg-12-encoding vectors. For example, *Saccharomyces cerevisiae* may be used. In addition a number of other genera, species, and strains are commonly available and useful herein, such as *Schizosaccharomyces pombe* (Beach *et al.* Nature, 290:140 (1981); EP 139,383);  
10 *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer *et al.*, *supra*) such as, *e.g.*, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt *et al.*, *J. Bacteriol.*, 737 (1983)), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg *et al.*, *supra*), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna *et al.* J. Basic Microbiol., 28:265-278 (1988));  
15 *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case *et al.* Proc. Natl. Acad. Sci. USA, 76:5259-5263 (1979)); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538); and filamentous fungi such as, *e.g.*, *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.* Biochem. Biophys. Res. Commun., 112:284-289 (1983); Tilburn *et al.*, *Gene*, 26:205-221 (1983); Yelton *et al.* Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) and *A. niger* (Kelly *et al.* EMBO J., 4:475-479 (1985)).

Transformation of appropriate cell hosts with a rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of  
25 vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed, see, for example, Cohen *et al.* Proc. Natl. Acad. Sci. USA 69:2110, (1972); and Maniatis *et al.*, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982). With regard to transformation of vertebrate cells with vectors  
30 containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for example, Graham *et al.* Virology, 52:456, (1973); Wigler *et al.* Proc.

Natl. Acad. Sci. USA 76:1373-76, (1979). The calcium phosphate precipitation method is preferred. However, other methods of for introducing DNA into cells may also be used, including nuclear microinjection and bacterial protoplast fusion.

5 For transient expression of recombinant channels, transformed host cells for the measurement of Na<sup>+</sup> current or intracellular Na<sup>+</sup> levels are typically prepared by co-transfecting constructs into cells such as HEK293 cells with a fluorescent reporter plasmid (such as pGreen Lantern-1, Life Technologies) using the calcium-phosphate precipitation technique (Ukomadu *et al.* Neuron 8, 663-676 (1992)). After forty-eight hours, cells with green fluorescence are selected for recording (Dib-Hajj *et al.* FEBS  
10 Lett. 416, 11-14 (1997)). Similarly, for transient expression of Mrg receptors and measurement of intracellular Ca<sup>2+</sup> changes in response to receptor activation as described in Example 4, HEK cells can be co-transfected with Mrg expression constructs and a fluorescent reporter plasmid. HEK293 cells are typically grown in high glucose DMEM (Life Technologies) supplemented with 10% fetal calf serum (Life  
15 Technologies).

Prokaryotic cells used to produce polypeptides of this invention are cultured in suitable media as described generally in Sambrook *et al.*, *supra*.

The mammalian host cells used to produce the polypeptides of this invention may be cultured in a variety of media, including but not limited to commercially  
20 available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma). In addition, any of the media described in Ham *et al.* Meth. Enz., 58:44 (1979), Barnes *et al.* Anal. Biochem. 102:255 (1980), U.S. Pat. Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; or 5,122,469; WO 90/03430; WO 87/00195; or U.S. Patent Re.  
25 30,985 may be used as culture media for the host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics, trace elements, and glucose or an equivalent energy source.  
30 Any other necessary supplements may also be included at appropriate concentrations as

determined by the skilled practitioner. The culture conditions are those previously used with the host cell selected for expression, and will be apparent to the skilled artisan.

The host cells referred to in this disclosure encompass cells in culture as well as cells that are within a host animal.

5           Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of  
10           the rDNA using a method such as that described by Southern, J. Mol. Biol. 98:503, (1975), or Berent *et al.*, Biotech. 3:208, (1985) or the proteins produced from the cell assayed via an immunological method as described below.

          Gene amplification and/or expression may be measured by any technique known in the art, including Southern blotting, Northern blotting to quantitate the transcription  
15           of mRNA (Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201-5205 (1980)), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly <sup>32</sup>P. Immunological methods for measuring gene expression include immunohistochemical staining of tissue sections or cells in culture, as well as  
20           assaying protein levels in culture medium or body fluids.. With immunohistochemical staining techniques, a cell sample is prepared by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like.

25           Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared as described herein.

## F. Production of Recombinant Proteins using an rDNA Molecule

The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

5           A nucleic acid molecule is first obtained that encodes a *mrg* or *drg-12* protein of the invention, for example, nucleotides 115-1026 of SEQ ID NO: 1, nucleotides 115-1029 of SEQ ID NO: 1, nucleotides 137-1051 of SEQ ID NO: 3, nucleotides 137-1054 of SEQ ID NO: 3, nucleotides 165-1070 of SEQ ID NO: 5, nucleotides 165-1073 of SEQ ID NO: 5, nucleotides 1-450 of SEQ ID NO: 7, nucleotides 1-459 of SEQ ID NO: 9, nucleotides 1820-2734 of SEQ ID NO: 11, nucleotides 170-574 of SEQ ID NO: 13, nucleotides 170-577 of SEQ ID NO: 13, nucleotides 328-1293 of SEQ ID NO: 15, nucleotides 328-1296 of SEQ ID NO: 15, nucleotides 171-1160 of SEQ ID NO: 17, nucleotides 171-1163 of SEQ ID NO: 17, nucleotides 83-943 of SEQ ID NO: 20, nucleotides 83-946 of SEQ ID NO: 20; nucleotides 16-918 of SEQ ID NO: 22, nucleotides 16-921 of SEQ ID NO: 22; nucleotides 106-1020 of SEQ ID NO: 24, nucleotides 106-1023 of SEQ ID NO: 24; nucleotides 45-959 of SEQ ID NO: 26, nucleotides 45-962 of SEQ ID NO: 26, nucleotides 1-405 of SEQ ID NO: 28 and nucleotides 1-408 of SEQ ID NO: 28. If the encoding sequence is uninterrupted by introns, as are these sequences, it is directly suitable for expression in any host.

20           The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated or when the recombinant cells are used, for instance, in high throughput assays.

25           Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth

above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

In one embodiment, Mrg or drg-12 may be produced by homologous recombination. Briefly, primary human cells containing an Mrg- or drg-12-encoding gene are transformed with a vector comprising an amplifiable gene (such as dihydrofolate reductase (DHFR)) and at least one flanking region of a length of at least about 150 bp that is homologous with a DNA sequence at the locus of the coding region of the Mrg or drg-12 gene. The amplifiable gene must be located such that it does not interfere with expression of the Mrg or drg-12 gene. Upon transformation the construct becomes homologously integrated into the genome of the primary cells to define an amplifiable region.

Transformed cells are then selected for by means of the amplifiable gene or another marker present in the construct. The presence of the marker gene establishes the presence and integration of the construct into the host genome. PCR, followed by sequencing or restriction fragment analysis may be used to confirm that homologous recombination occurred.

The entire amplifiable region is then isolated from the identified primary cells and transformed into host cells. Clones are then selected that contain the amplifiable region, which is then amplified by treatment with an amplifying agent. Finally, the host cells are grown so as to express the gene and produce the desired protein.

The proteins of this invention may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide. In one embodiment the heterologous polypeptide may be a signal sequence. In general, the signal sequence may be a component of the vector, or it may be a part of the Mrg or drg-12 DNA that is inserted into the vector. The heterologous signal sequence selected preferably is one that is recognized and processed (*i.e.*, cleaved by a signal peptidase) by the host cell. For expression in prokaryotic host cells the signal sequence may be a prokaryotic signal

sequence selected, for example, from the group consisting of the alkaline phosphatase, penicillinase, lpp, and heat-stable enterotoxin II leaders. For yeast secretion the native signal sequence may be substituted by, e.g., the yeast invertase leader,  $\alpha$  factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, or acid phosphatase leader and the *C. albicans* glucoamylase leader). In mammalian cell expression any native signal sequence is satisfactory. Alternatively it may be substituted with a signal sequence from related proteins, as well as viral secretory leaders, for example, the herpes simplex gD signal. The DNA for such precursor regions is ligated in reading frame to DNA encoding the mature protein or a soluble variant thereof.

The heterologous polypeptide may also be a marker polypeptide that can be used, for example, to identify the location of expression of the fusion protein. The marker polypeptide may be any known in the art, such as a fluorescent protein. A preferred marker protein is green fluorescent protein (GFP).

#### **G. Modifications of Mrg polypeptides**

Covalent modifications of Mrg and drg-12 and their respective variants are included within the scope of this invention. In one embodiment, specific amino acid residues of a polypeptide of the invention are reacted with an organic derivatizing agent. Derivatization with bifunctional agents is useful, for instance, for crosslinking Mrg or Mrg fragments or derivatives to a water-insoluble support matrix or surface for use in methods for purifying anti-Mrg antibodies and identifying binding partners and ligands. In addition, Mrg or Mrg fragments may be crosslinked to each other to modulate binding specificity and effector function. Many crosslinking agents are known in the art and include, but are not limited to, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other contemplated modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, Proteins: Structure and Molecular Properties, W.H.

Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Modification of the glycosylation patterns of the polypeptides of the invention are also contemplated. Methods for altering the glycosylation pattern of polypeptides are well known in the art. For example, one or more of the carbohydrate moities found in native sequence Mrg or drg-12 may be removed chemically, enzymatically or by modifying the glycosylation site. Alternatively, additional glycosylation can be added, such as by manipulating the composition of the carbohydrate moities directly or by adding glycosylation sites not present in the native sequence Mrg or drg-12 by altering the amino acid sequence.

Another type of covalent modification of the polypeptides of the invention comprises linking the polypeptide or a fragment or derivative thereof to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The polypeptides of the present invention may also be modified in a way to form a chimeric molecule comprising Mrg or drg-12 fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the Mrg or drg-12 with a tag polypeptide that provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the polypeptide. The epitope tag allows for identification of the chimeric protein as well as purification of the chimeric protein by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. A number of tag polypeptides and their respective antibodies are well known in the art. Well known tags include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flue HA tag polypeptide (Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)); the c-myc tag (Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)); the Herpes Simplex virus glycoprotein D (gD) tag (Paborsky et al., Protein Engineering, 3(6):547-553 (1990)) and the Flag-peptide (Hopp et al., BioTechnology, 6:1204-1210 (1988)).

In another embodiment, the chimeric molecule comprises a fusion of Mrg or drg-12 with an immunoglobulin or a particular region of an immunoglobulin. To produce an immunoadhesin, the polypeptide of the invention or a fragment or specific domain(s) thereof could be fused to the Fc region of an IgG molecule. Typically the fusion is to an immunoglobulin heavy chain constant region sequence. Mrg- or drg-12-immunoglobulin chimeras for use in the present invention are normally prepared from nucleic acid encoding one or more extracellular domains, or fragments thereof, of an Mrg or drg-12 receptor fused C-terminally to nucleic acid encoding the N-terminus of an immunoglobulin constant domain sequence. N-terminal fusions are also possible.

While not required in the immunoadhesins of the present invention, an immunoglobulin light chain might be present either covalently linked to an Mrg- or drg-12-immunoglobulin heavy chain fusion polypeptide, or directly fused to Mrg or drg-12. In order to obtain covalent association, DNA encoding an immunoglobulin light chain may be coexpressed with the DNA encoding the Mrg- or drg-12-immunoglobulin heavy chain fusion protein. Upon secretion, the hybrid heavy chain and the light chain will be covalently associated to provide an immunoglobulin-like structure comprising two disulfide-linked immunoglobulin heavy chain-light chain pairs.

Bispecific immunoadhesins may also be made. Such immunoadhesins may combine an Mrg or drg-12 domain and a domain, such as the extracellular domain, from another receptor. Alternatively, the immunoadhesins herein might comprise portions of two different Mrg receptors, each fused to an immunoglobulin heavy chain constant domain sequence.

In yet another embodiment, the chimeric molecule of the present invention comprises a fusion of Mrg or drg-12 or a fragment or domain(s) thereof, with a heterologous receptor or fragment or domain(s) thereof. The heterologous receptor may be a related Mrg or drg-12 family member, or may be completely unrelated. The heterologous protein fused to the Mrg or drg-12 protein may be chosen to obtain a fusion protein with a desired ligand specificity or a desired affinity for a particular ligand or to obtain a fusion protein with a desired effector function.



## H. Methods of Using mrgs or drgs as Molecular or Diagnostic Probes

5 The sequences and antibodies, proteins and peptides of the present invention may be used as molecular probes for the detection of cells or tissues related to or involved with sensory perception, especially perception of pain. Although many methods may be used to detect the nucleic acids or proteins of the invention in situ, preferred probes include antisense molecules and anti-mrg or anti-drg-12 antibodies.

10 Probes for the detection of the nucleic acids or proteins of the invention may find use in the identification of the involvement of Mrg or drg-12 proteins in particular disease states, such as glaucoma or chronic pain, or in enhanced or inhibited sensory perception. In particular, probes of the present invention may be useful in determining if Mrg or drg-12 expression is increased or decreased in patients demonstrating changes in sensory perception, such as in patients with allodynia, hyperalgesia or chronic pain, or patients with a disease or disorder, such as glaucoma. A determination of decreased expression or overexpression of a polypeptide of the invention may be useful in  
15 identifying a therapeutic approach to treating the disorder, such as by administering Mrg or drg-12 agonists or antagonists.

Determination of changes in Mrg or drg-12 expression levels in animal models of disease states, particularly pain, may also be useful in identifying the types of disorders that might be effectively treated by compounds that modify expression or  
20 activity.

Further, the probes of the invention, including antisense molecules and antibodies, may be used to detect the expression of mutant or variant forms of Mrg or drg-12 variants. The ability to detect such variants may be useful in identifying the role that the variants play in particular disease states and in the symptoms experienced by  
25 particular patients. Identification of the involvement of a variant of Mrg or drg-12 in a disease or disorder may suggest a therapeutic approach for treatment of the disease or disorder, such as gene therapy or the administration of agonists or antagonists known to bind the receptor variant.

In addition, probes of the invention may be used to determine the exact  
30 expression patterns of the various Mrg and drg-12 family members, including the relationship of one to another. For example, the microscopy images of in situ

hybridization in Figure 2 show the localization of antisense staining against a nucleotide of SEQ ID NO:2 (“*mrg3*”) and of SEQ ID NO:4 (“*mrg4*”) in transverse sections of dorsal root ganglia (DRG) from newborn wild type (WT) and *Neurogenin1* null mutant (*Ngn1*<sup>-/-</sup>) mice. White dashed lines outline the DRG and black dashed lines outline the spinal cord. Note that in the *Ngn1*<sup>-/-</sup> mutant, the size of the DRG is severely reduced due to the loss of nociceptive sensory neurons, identified using three other independent markers (*trkA*; VR-1 and SNS-TTXi (Ma et al., (1999)). *mrg3* is expressed in a subset of DRG in WT mice (A) but is absent in the *Ngn1*<sup>-/-</sup> DRG (B). *mrg4* is expressed in a smaller subset of DRG than that of *mrg3* (C). It is also absent in the *Ngn1*<sup>-/-</sup> DRG (D). The loss of *mrg*-expressing neurons in the *Ngn1*<sup>-/-</sup> DRG indicates that these neurons are likely to be nociceptive.

Expression of *mrgs* in subsets of dorsal root ganglia (DRG) neurons are shown in Figure 2A. Frozen transverse sections of DRG from wild-type (a-i) and *ngn1*<sup>-/-</sup> (j) mutant new born mice were annealed with antisense digoxigenin RNA probes, and hybridization was visualized with an alkaline phosphatase-conjugated antibody. Positive signals are shown as dark purple stainings. *TrkA* is expressed in a large portion of wild-type DRG neurons (a) but absent in *ngn1*<sup>-/-</sup> (data not shown). Each of the eight *mrg* genes (b-i) is expressed in a small subset of neurons in wild-type DRG in completely absent in *ngn1*<sup>-/-</sup> DRG (j and data not shown). Black dash line outlines the *ngn1*<sup>-/-</sup> mutant DRG.

In Figure 2B, *mrgs* are expressed by *TrkA*<sup>+</sup> nociceptive neurons. Double labeling technique was used to colocalize *TrkA* (green; [b,e]) and *mrgs* (red; [a,d]) in DRG neurons. During the double labeling experiments frozen sections of wild-type DRG were undergone in situ hybridizations with either *mrg3* (a-c) or *mrg5* (d-f) fluorescein-labeled antisense RNA probes followed by anti-*TrkA* antibody immunostaining. The same two frames (a and b, d and e) were digitally superimposed to reveal the extent of colocalization (c, f). The colocalizations of *TrkA* with either *mrg3* or *mrg5* appear yellow in merged images (c, f, respectively). The white arrowheads indicate examples of double positive cells.

In Figure 2C, *mrgs* and VR1 define two different populations of nociceptive neurons in DRG. The combination of in situ hybridizations (red) with either *mrg3* or

*mrg5* fluorescein-labeled antisense RNA probes and anti-VR1 antibody immunostaining (green) demonstrated that neither *mrg3* (a-c) nor *mrg5* (d-f) were expressed by VR1-positive neurons. In the merged images (c,f), there are no colocalizations of VR1 with either *mrg3* or *mrg5*. The white arrowheads are pointed to *mrgs*-expressing but VR1-negative nociceptive neurons.

In Figure 2D *mrgs* are shown to be expressed by IB4<sup>+</sup> nociceptive neurons. Double labeling technique was used to colocalize IB4 (green; [b,e]) and *mrgs* (red; [a,d]) in DRG neurons. The expressions of *mrg3* and *mrg5* were visualized by in situ hybridization as described before. The same DRG sections were subsequently undergone through FITC-conjugated lectin IB4 binding. In the merged images (c,f), there are extensive overlappings between *mrgs* and IB4 stainings (yellow neurons indicated by arrowheads).

Information about the expression patterns of the receptors of the invention in normal tissue and tissue taken from animal models of disease or patients suffering from a disease or disorder will be useful in further defining the biological function of the receptors and in tailoring treatment regimens to the specific receptor or combination of receptors involved in a particular disease or disorder.

## I. Methods to Identify Binding Partners

As discussed in more detail below, several peptides have been putatively identified as endogenous ligands for Mrg receptors. In particular the RF-amide peptides, including NPAF and NPFF, have been shown to efficiently stimulate several of the Mrg receptors. In order to identify additional new ligands for the Mrg receptors and ligands for drg-12, it is first necessary to identify compounds that bind to these receptors. Thus, another embodiment of the present invention provides methods of isolating and identifying binding partners or ligands of proteins of the invention. Macromolecules that interact with Mrg are referred to, for purposes of this discussion, as "binding partners." While the discussion below is specifically directed to identifying binding partners for Mrg receptors, it is contemplated that the assays of the invention may be used to identify binding partners for drg-12 as well.

Receptor binding can be tested using Mrg receptors isolated from their native source or synthesized directly. However, Mrg receptors obtained by the recombinant methods described above are preferred.

The compounds which may be screened in accordance with the invention include, but are not limited to polypeptides, peptides, including but not limited to members of random peptide libraries; (see, *e.g.*, Lam, K.S. *et al.*, 1991, *Nature* 354:82-84; Houghten, R. *et al.*, 1991, *Nature* 354:84-86) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids, phosphopeptides (including, but not limited to members of random or partially degenerate, directed phosphopeptide libraries; see, *e.g.*, Songyang, Z. *et al.*, 1993, *Cell* 72:767-778), peptide mimetics, antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, FAb, F(abN)<sub>2</sub> and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

The ability of candidate or test compounds to bind Mrg receptors can be measured directly or indirectly, such as in competitive binding assays. In competitive binding experiments, the concentration of the test compound necessary to displace 50% of another compound bound to the receptor (IC<sub>50</sub>) is used as a measure of binding affinity. In these experiments the other compound is a ligand known to bind to the Mrg receptor with high affinity, such as an RF-amide peptide.

A variety of assay formats may be employed, including biochemical screening assays, immunoassays, cell-based assays and protein-protein binding assays, all of which are well characterized in the art. In one embodiment the assay involves anchoring the test compound onto a solid phase, adding the non-immobilized component comprising the Mrg receptor, and detecting Mrg/test compound complexes anchored on the solid phase at the end of the reaction. In an alternative embodiment, the Mrg may be anchored onto a solid surface, and the test compound, which is not anchored. In both situations either the test compound or the Mrg receptor is labeled, either directly or indirectly, to allow for identification of complexes. For example, an Mrg-Ig immunoadhesin may be anchored to a solid support and contacted with one or more test compounds.

Microtiter plates are preferably utilized as the solid phase and the anchored component may be immobilized by non-covalent or covalent attachments. Non-covalent attachment may be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized may be used to anchor the protein to the solid surface.

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; *e.g.*, using an immobilized antibody specific for either Mrg polypeptide, peptide or fusion protein or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

In one embodiment of these methods, a protein of the invention or a fragment of a protein of the invention, for instance, an extracellular domain fragment, is mixed with one or more potential binding partners, or an extract or fraction of a cell, under conditions that allow the association of potential binding partners with the protein of the invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be removed, identified and further analyzed. To identify and isolate a binding partner, the entire Mrg protein, for instance a protein comprising the entire amino acid sequence of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 16, 18, 21, 23, 25, 27, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 or 109 can be used. Alternatively, a fragment of the Mrg polypeptide can be used.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from DRG. Alternatively, cellular extracts may be prepared from cells derived from any tissue, including normal human kidney tissue, or available cell lines, particularly kidney derived cell lines.

A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical

disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

5           Once an extract of a cell is prepared, the extract is mixed with the protein of the invention under conditions in which association of the protein with the binding partner can occur. Alternatively, one or more known compounds or molecules can be mixed with the protein of the invention. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a  
10   human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

          After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For  
15   example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

          After removal of non-associated cellular constituents found in the extract, and/or  
20   unbound compounds or molecules, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

          To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the  
25   protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins or any other macromolecule.

          Alternatively, binding partners may be identified using a Far-Western assay  
30   according to the procedures of Takayama *et al.* Methods Mol. Biol. 69:171-84 (1997) or

Sauder *et al.* J Gen.Virol. 77(5): 991-6 or identified through the use of epitope tagged proteins or GST fusion proteins.

5 Binding partners may also be identified in whole cell binding assays that are well known in the art. In one embodiment, an Mrg receptor is expressed in cells in which it is not normally expressed, such as COS cells. The cells expressing Mrg are then contacted with a potential binding partner that has previously been labeled, preferably with radioactivity or a fluorescent marker. The cells are then washed to remove unbound material and the binding of the potential binding partner to the cells is assessed, for example by collecting the cells on a filter and counting radioactivity. The amount of binding of the potential binding partner to untransfected cells or mock transfected cells is subtracted as background.

15 This type of assay may be carried out in several alternative ways. For example, in one embodiment it is done using cell membrane fractions from cells transfected with an Mrg or known to express an Mrg, rather than whole cells. In another embodiment purified Mrg is refolded in lipids to produce membranes that are used in the assay.

Alternatively, the nucleic acid molecules of the invention can be used in cell based systems to detect protein-protein interactions (see WO99/55356). These systems have been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

20 Any method suitable for detecting protein-protein interactions may be employed for identifying proteins, including but not limited to soluble, transmembrane or intracellular proteins, that interact with Mrg receptors. Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns to identify proteins that interact with Mrg. For such assays, the Mrg component can be a full-length protein, a soluble derivative thereof, a peptide corresponding to a domain of interest, or a fusion protein containing some region of Mrg.

25 Methods may be employed which result in the simultaneous identification of genes that encode proteins capable of interacting with Mrg. These methods include, for example, probing expression libraries, using labeled Mrg or a variant thereof.

One method of detecting protein interactions *in vivo* that may be used to identify Mrg binding partners is the yeast two-hybrid system. This system is well known in the art and is commercially available from Clontech (Palo Alto, CA).

5 Briefly, two hybrid proteins are employed, one comprising the DNA-binding domain of a transcription activator protein fused to the Mrg receptor, or a polypeptide, peptide, or fusion protein therefrom, and the other comprising the transcription activator protein's activation domain fused to an unknown target protein. These proteins are expressed in a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (e.g., HBS or *lacZ*) whose regulatory region contains the transcription activator's  
10 binding site. While either hybrid protein alone cannot activate transcription of the reporter gene, interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The target protein is preferably obtained from tissue or cells known to express  
15 the Mrg receptor, such as DRG cells. For example, a cDNA library prepared from DRG cells may be used.

Binding partners may also be identified by their ability to interfere with or disrupt the interaction of known ligands. Even if they do not activate Mrg receptors, binding partners that interfere with interactions with known ligands may be useful in  
20 regulating or augmenting Mrg activity in the body and/or controlling disorders associated with Mrg activity (or a deficiency thereof).

Compounds that interfere with the interaction between Mrg and a known ligand may be identified by preparing a reaction mixture containing Mrg, or some variant or fragment thereof, and a known binding partner, such as an RF-amide peptide, under  
25 conditions and for a time sufficient to allow the two to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound may be initially included in the reaction mixture, or may be added at a time subsequent to the addition of the Mrg and its binding partner. Control reaction mixtures are incubated  
30 without the test compound. The formation of any complexes between the Mrg and the binding partner is then detected. The formation of a complex in the control reaction, but



not in the reaction mixture containing the test compound indicates that the compound interferes with the interaction of the Mrg and the known binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal Mrg protein may also be compared to complex formation within reaction mixtures containing the test compound and a mutant Mrg. This comparison may be important in those cases wherein it is desirable to identify compounds that specifically disrupt interactions of mutant, or mutated, Mrg but not the normal proteins.

The order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction by competition can be identified by conducting the binding reaction in the presence of the test substance. In this case the test compound is added to the reaction mixture prior to, or simultaneously with, Mrg and the known binding partner. Alternatively, test compounds that have the ability to disrupt preformed complexes can be identified by adding the test compound to the reaction mixture after complexes have been formed.

In an alternate embodiment of the invention, a preformed complex of Mrg and an interactive binding partner is prepared in which either the Mrg or its binding partners is labeled, but the signal generated by the label is quenched due to formation of the complex (see, *e.g.*, U.S. Patent No. 4,109,496 to Rubenstein which utilizes this approach for immunoassays). The addition of a test compound that competes with and displaces one of the species from the preformed complex results in the generation of a signal above background. In this way, test substances which disrupt the interaction can be identified.

Whole cells expressing Mrg, membrane fractions prepared from cells expressing Mrg or membranes containing refolded Mrg may be used in the assays described above. However, these same assays can be employed using peptide fragments that correspond to the binding domains of Mrg and/or the interactive or binding partner (in cases where the binding partner is a protein), in place of one or both of the full length proteins. Any number of methods routinely practiced in the art can be used to identify and isolate the binding sites. These methods include, but are not limited to, mutagenesis of the gene encoding an Mrg protein and screening for disruption of binding of a known ligand.

The compounds identified can be useful, for example, in modulating the activity of wild type and/or mutant Mrg; can be useful in elaborating the biological function of Mrg receptors; can be utilized in screens for identifying compounds that disrupt normal Mrg receptor interactions or may themselves disrupt or activate such interactions; and can be useful therapeutically.

#### J. Methods to Identify Agents that Modulate the Expression of a Nucleic Acid.

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding a *mrg* or *drg-12* protein of the invention or another protein involved in an *mrg* or *drg-12* mediated pathway. These agents may be, but are not limited to, peptides, peptide mimetics, and small organic molecules that are able to gain entry into an appropriate cell (e.g., in the DRG) and affect the expression of a gene. Agents that modulate the expression of Mrg or drg-12 or a protein in an *mrg* mediated pathway may be useful therapeutically, for example to increase or decrease sensory perception, such as the perception of pain, to treat glaucoma, or to increase or decrease wound healing.

Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding the protein having the sequence of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107 or 109 if it is capable of up- or down-regulating expression of the gene or mRNA levels nucleic acid in a cell.

In one assay format, cell lines that contain reporter gene fusions between the open reading frames and/or the 5' or 3' regulatory sequences of a gene of the invention and any assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al. Anal. Biochem.* 188:245-254 (1990)). Cell lines containing the reporter gene fusions are then exposed to the agent to

be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding a *mrg* or *drg-12* protein.

5 Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding a *mrg* or *drg-12* protein of the invention. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.* (*Molecular Cloning: A Laboratory*  
10 *Manual*, 2nd Ed. Cold Spring Harbor Laboratory Press, 1989).

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid  
15 hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

20 Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Laboratory Press, NY, 1989) or Ausubel *et al.* (*Current*  
25 *Protocols in Molecular Biology*, Greene Publishing Co., NY, 1995).

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.* and Ausubel *et al.*, as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA  
30 enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the

invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a silicon chip or porous glass wafer. The wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding a *mrg* or *drg-12* are identified.

Hybridization for qualitative and quantitative analysis of mRNAs may also be carried out by using a RNase Protection Assay (*i.e.*, RPA, see Ma *et al.* Methods 10: 273-238 (1996)). Briefly, an expression vehicle comprising cDNA encoding the gene product and a phage specific DNA dependent RNA polymerase promoter (*e.g.*, T7, T3 or SP6 RNA polymerase) is linearized at the 3' end of the cDNA molecule, downstream from the phage promoter, wherein such a linearized molecule is subsequently used as a template for synthesis of a labeled antisense transcript of the cDNA by *in vitro* transcription. The labeled transcript is then hybridized to a mixture of isolated RNA (*i.e.*, total or fractionated mRNA) by incubation at 45°C overnight in a buffer comprising 80% formamide, 40 mM Pipes, pH 6.4, 0.4 M NaCl and 1 mM EDTA. The resulting hybrids are then digested in a buffer comprising 40 µg/ml ribonuclease A and 2 µg/ml ribonuclease. After deactivation and extraction of extraneous proteins, the samples are loaded onto urea/polyacrylamide gels for analysis.

In another assay format, products, cells or cell lines are first be identified which express *mrg* or *drg-12* gene products physiologically. Cells and/or cell lines so identified would be expected to comprise the necessary cellular machinery such that the fidelity of modulation of the transcriptional apparatus is maintained with regard to exogenous contact of agent with appropriate surface transduction mechanisms and/or the cytosolic cascades. Such cells or cell lines are then transduced or transfected with an expression vehicle (*e.g.*, a plasmid or viral vector) construct comprising an operable non-translated 5' or 3'-promoter containing end of the structural gene encoding the

instant gene products fused to one or more antigenic fragments, which are peculiar to the instant gene products, wherein said fragments are under the transcriptional control of said promoter and are expressed as polypeptides whose molecular weight can be distinguished from the naturally occurring polypeptides or may further comprise an immunologically distinct tag. Such a process is well known in the art.

Cells or cell lines transduced or transfected as outlined above are then contacted with agents under appropriate conditions; for example, the agent comprises a pharmaceutically acceptable excipient and is contacted with cells comprised in an aqueous physiological buffer such as phosphate buffered saline (PBS) at physiological pH, Eagles balanced salt solution (BSS) at physiological pH, PBS or BSS comprising serum or conditioned media comprising PBS or BSS and/or serum incubated at 37° C. Said conditions may be modulated as deemed necessary by one of skill in the art. Subsequent to contacting the cells with the agent, said cells will be disrupted and the polypeptides of the lysate are fractionated such that a polypeptide fraction is pooled and contacted with an antibody to be further processed by immunological assay (*e.g.*, ELISA, immunoprecipitation or Western blot). The pool of proteins isolated from the “agent-contacted” sample will be compared with a control sample where only the excipient is contacted with the cells and an increase or decrease in the immunologically generated signal from the “agent-contacted” sample compared to the control will be used to distinguish the effectiveness of the agent.

The probes described above for identifying differential expression of Mrg mRNA in response to applied agents can also be used to identify differential expression of Mrg mRNA in populations of mammals, for example populations with differing levels of sensory perception. Methods for identifying differential expression of genes are well known in the art. In one embodiment, mRNA is prepared from tissue or cells taken from patients exhibiting altered sensory perception, such as patients experiencing neuropathic pain, or suffering from a disease or disorder in which the Mrg receptor may play a role, such as glaucoma, and Mrg expression levels are quantified using the probes described above. The Mrg expression levels may then be compared to those in other populations to determine the role that Mrg expression is playing in the alteration of

sensory perception and to determine whether treatment aimed at increasing or decreasing Mrg expression levels would be appropriate.

**K. Methods to Identify Agents that Modulate Protein Levels or at Least One Activity of the Proteins of DRG Primary Sensory Neurons.**

Another embodiment of the present invention provides methods for identifying agents or conditions that modulate protein levels and/or at least one activity of a *mrg* or *drg-12* protein of the invention, including agonists and antagonists. Such methods or assays may utilize any means of monitoring or detecting the desired activity.

In one format, the relative amounts of a protein of the invention between a cell population that has been exposed to the agent to be tested compared to an unexposed control cell population may be assayed. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe.

In another embodiment, animals known to express Mrg or drg-12 receptors are subjected to a particular environmental stimulus and any change produced in Mrg or drg-12 protein expression by exposure to the stimulus is measured. Transgenic animals, such as transgenic mice, produced to express a particular Mrg in a particular location may be used. The environmental stimulus is not limited and may be, for example, exposure to stressful conditions, or exposure to noxious or painful stimuli. Differences in Mrg receptor expression levels in response to environmental stimuli may provide insight into the biological role of Mrgs and possible treatments for diseases or disorders related to the stimuli used.

Antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins of the invention if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic

conjugates with carriers such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co. (Rockford, IL), may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein Nature 256:495-497 (1975)) or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonals or the polyclonal antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as the Fab, Fab', of F(ab')<sub>2</sub> fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin, such as humanized antibodies as discussed in more detail below.

## 1. Identification of Agonists and Antagonists

The present invention provides for assays to identify compounds that serve as agonists or antagonists of one or more of the biological properties of Mrg and/or *drg-12*. Mrg agonists and antagonists may be useful in the prevention and treatment of problems associated with sensory perception, particularly nociception. For example, compounds identified as Mrg receptor agonists may be used to stimulate Mrg receptor activation and thus may be effective in treating mammals suffering from pain. Compounds that are identified as Mrg receptor antagonists may be used, for example, to decrease the effector functions of Mrg receptors. This may be useful in cases where the Mrg receptors contain a mutation that produces increased responsiveness, or in cases of Mrg receptor overexpression. For instance, Mrg receptor antagonists may be useful in increasing the sensitivity of mammals to pain where appropriate, such as in diseases involving decreased sensory responsiveness, like some forms of diabetes.

Assays for identifying agonists or antagonists may be done *in vitro* or *in vivo*, by monitoring the response of a cell following binding of the ligand to the receptor. An agonist will produce a cellular response, while an antagonist will have no effect on cellular response but will be capable of preventing cellular response to a known agonist.

### a. Small Molecules

Small molecules may have the ability to act as Mrg agonists or antagonists and thus may be screened for an effect on a biological activity of Mrg. Small molecules preferably have a molecular weight of less than 10 kD, more preferably less than 5 kD and even more preferably less than 2 kD. Such small molecules may include naturally occurring small molecules, synthetic organic or inorganic compounds, peptides and peptide mimetics. However, small molecules in the present invention are not limited to these forms. Extensive libraries of small molecules are commercially available and a wide variety of assays are well known in the art to screen these molecules for the desired activity.

Candidate Mrg agonist and antagonist small molecules are preferably first identified in an assay that allows for the rapid identification of potential agonists and antagonists. An example of such an assay is a binding assay wherein the ability of the candidate molecule to bind to the Mrg receptor is measured, such as those described



above. In another example, the ability of candidate molecules to interfere with the binding of a known ligand, for example FMRFamide to MrgA1, is measured. Candidate molecules that are identified by their ability to bind to Mrg proteins or interfere with the binding of known ligands are then tested for their ability to stimulate one or more biological activities.

The activity of the proteins of the invention may be monitored in cells expressing the *mrg* and/or *drg-12* proteins of the invention by assaying for physiological changes in the cells upon exposure to the agent or agents to be tested. Such physiological changes include but are not limited to the flow of current across the membrane of the cell.

In one embodiment the protein is expressed in a cell that is capable of producing a second messenger response and that does not normally express Mrg or *drg-12*. The cell is then contacted with the compound of interest and changes in the second messenger response are measured. Methods to monitor or assay these changes are readily available. For instance, the *mrg* genes of the invention may be expressed in cells expressing Gα15, a G protein α subunit that links receptor activation to increases in intracellular calcium [Ca<sup>2+</sup>] which can be monitored at the single cell level using the FURA-2 calcium indicator dye as disclosed in Chandrashekar *et al.* Cell 100:703-711, (2000). This assay is described in more detail in Example 5.

Similar assays may also be used to identify inhibitors or antagonists of Mrg or *drg-12* activation. For example, cells expressing Mrg or *drg-12* and capable of producing a quantifiable response to receptor activation are contacted with a known Mrg or *drg-12* activator and the compound to be tested. In one embodiment, HEK cells expressing Gα15 and MrgA1 are contacted with FMRFamide and the compound to be tested. The cellular response is measured, in this case increase in [Ca<sup>2+</sup>]. A decreased response compared to the known activator by itself indicates that the compound acts as an inhibitor of activation.

While such assays may be formatted in any manner, particularly preferred formats are those that allow high -throughput screening (HTP). In HTP assays of the invention, it is possible to screen thousands of different modulators or ligands in a single day. For instance, each well of a microtiter plate can be used to run a separate

assay, for instance an assay based on the ability of the test compounds to modulate receptor activation derived increases in intracellular calcium as described above.

Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial library, or a growth broth of an organism.

As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a nonrandom basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. Sites of interest might be peptides within the membrane spanning regions, cytoplasmic and extracellular peptide loops between these transmembrane regions, or selected sequences within the N-terminal extracellular domain or C-terminal intracellular domain. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. Dominant negative proteins, DNAs encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be introduced into cells to affect function. "Mimic" used herein refers to the modification of a region or several regions of a peptide molecule to provide a structure chemically different from the parent peptide but topographically and functionally similar to the parent peptide (see Grant GA. in: Meyers (ed.) Molecular Biology and Biotechnology (New York, VCH Publishers, 1995), pp. 659-664). A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard

recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

#### b. Antibodies

Another class of agents of the present invention are antibodies immunoreactive with critical positions of proteins of the invention. These antibodies may be human or non-human, polyclonal or monoclonal and may serve as agonist antibodies or neutralizing antibodies. They include amino acid sequence variants, glycosylation variants and fragments of antibodies. Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies. General techniques for the production of such antibodies and the selection of agonist or neutralizing antibodies are well known in the art.

The antibodies of the present invention can be polyclonal antibodies, monoclonal antibodies, chimeric antibodies, humanized antibodies, human antibodies, heteroconjugate antibodies, or antibody fragments. In addition, the antibodies can be made by any method known in the art, including recombinant methods.

Mrg agonist and neutralizing antibodies may be preliminarily identified based on their ability to bind the Mrg receptor. For example, Western blot techniques well known in the art may be used to screen a variety of antibodies for their ability to bind Mrg. Mrg agonist and neutralizing antibodies are then identified from the group of candidate antibodies based on their biological activity. In one embodiment, Mrg agonist antibodies are identified by their ability to induce activation of a second messenger system in cells expressing the Mrg protein and comprising a second messenger system. For example, HEK cells overexpressing Gα15 and transfected with *mrg* may be contacted with a potential Mrg agonist antibody. An increase in intracellular calcium, measured as described in Example 5, would indicate that the antibody is an agonist antibody.

Identification of a neutralizing antibody involves contacting a cell expressing Mrg with a known Mrg ligand, such as an RF-amide peptide, and the candidate antibody and observing the effect of the antibody on Mrg activation. In one embodiment, Mrg receptors expressed in HEK cells overexpressing Gα15 are contacted

with an Mrg ligand such as FMRFamide and the candidate neutralizing antibody. A decrease in responsiveness to the ligand, measured as described in Example 5, would indicate that the antibody is a neutralizing antibody.

5                   c. Other antagonists

The Mrg or drg-12 antagonists are not limited to Mrg or drg-12 ligands. Other antagonists include variants of a native Mrg or drg-12 receptor that retains the ability to bind an endogenous ligand but is not able to mediate a biological response. Soluble receptors and immunoadhesins that bind Mrg or drg-12 ligands may also be antagonists,  
10 as may antibodies that specifically bind a ligand near its binding site and prevent its interaction with the native receptor. These antagonists may be identified in the assays described above.

                  d. Computer Modeling

15           Computer modeling and searching technologies permit identification of compounds, or the improvement of already identified compounds, that can modulate Mrg receptor expression or activity. Once an agonist or antagonist is identified, the active sites or regions, such as ligand binding sites, are determined. The active site can be identified using methods known in the art including, for example, by determining the  
20 effect of various amino acid substitutions or deletions on ligand binding or from study of complexes of the relevant compound or composition with its natural ligand, such as with X-ray crystallography.

Next, the three dimensional geometric structure of the active site is determined such as by X-ray crystallography, NMR, chemical crosslinking or other methods known  
25 in the art. Computer modeling can be utilized to make predictions about the structure where the experimental results are not clear. Examples of molecular modeling systems are the CHARMM and QUANTA programs (Polygen Corporation, Waltham, MA). Once a predicted structure is determined, candidate modulating compounds can be identified by searching databases containing compounds along with information on their  
30 molecular structure in an effort to find compounds that have structures capable of interacting with the active site. The compounds found from this search are potential

modulators of the activity of the proteins of the present invention and can be tested in the assays described above.

The agonistic or antagonistic activity of test compounds identified in cell based assays as described above can be further elucidated in assays using animals, for example transgenic animals that overexpress Mrg receptors as described in more detail below. In one embodiment, the effect of administration of potential Mrg antagonists or agonists on the responsiveness of such transgenic animals to sensory stimuli, such as noxious or painful stimuli, is measured. The therapeutic utility of such compounds may be confirmed by testing in these types of experiments or in animal models of particular disorders, for example animal models of neuropathic pain.

#### **L. Uses for Agents that modulate at Least One Activity of the Proteins.**

As provided in the Examples, the *mrg* or *drg-12* proteins and nucleic acids of the invention, are expressed in the primary nociceptive sensory neurons of DRG. In addition the Mrg receptors are expressed in specialized skin cells that play a role in wound repair. Further, proteins homologous to Mrg receptors are expressed in the trabecular meshwork of the eye and a role for them has been suggested in the regulation of pressure in the eye (Gonzalez et al. Invest. Ophth. Vis. Sci. 41: 3678-3693 (2000)). Thus, the present invention further provides compositions containing one or more agents that modulate expression or at least one activity of a protein of the invention. For example, the invention provides ligands that directly activate Mrg receptors.

Agents that modulate, up-or-down-regulate the expression of the protein or agents such as agonists or antagonists of at least one activity of the protein may be used to modulate biological and pathologic processes associated with the protein's function and activity. Several agents that activate the Mrg receptors are identified in the examples, including the RF-amide peptides. Thus the present invention provides methods to treat impaired sensory perception, such as pain, including neuropathic pain, as well as to promote wound healing, to restore normal sensitivity following injury and to treat ocular conditions, particularly those associated with pressure, such as glaucoma.

As described in the Figures and Examples, expression of a protein of the invention may be associated with biological processes of nociception, which may also be considered pathological processes. As used herein, an agent is said to modulate a biological or pathological process when the agent alters the degree, severity or nature of the process. For instance, the neuronal transmission of pain signals may be prevented or modulated by the administration of agents which up-regulate down-regulate or modulate in some way the expression or at least one activity of a protein of the invention.

The pain that may be treated by the proteins of the present invention and agonists and antagonists thereof, is not limited in any way and includes pain associated with a disease or disorder, pain associated with tissue damage, pain associated with inflammation, pain associated with noxious stimuli of any kind, and neuropathic pain, including pain associated with peripheral neuropathies, as well as pain without an identifiable source. The pain may be subjective and does not have to be associated with an objectively quantifiable behavior or response.

In addition to treating pain, the compounds and methods of the present invention may be useful for increasing or decreasing sensory responses. It may be useful to increase responsiveness to stimuli, including noxious stimuli and painful stimuli, in some disease states that are characterized by a decreased responsiveness to stimuli, for example in diabetes.

Certain conditions, such as chronic disease states associated with pain and peripheral neuropathies and particularly conditions resulting from a defective Mrg gene, can benefit from an increase in the responsiveness to Mrg receptor ligands. Thus these condition may be treated by increasing the number of functional Mrg receptors in cells of patients suffering from such conditions. This could be increasing the expression of Mrg receptor in cells through gene therapy using Mrg-encoding nucleic acid. This includes both gene therapy where a lasting effect is achieved by a single treatment, and gene therapy where the increased expression is transient. Selective expression of Mrg in appropriate cells may be achieved by using Mrg genes controlled by tissue specific or inducible promoters or by producing localized infection with replication defective viruses carrying a recombinant Mrg gene, or by any other method known in the art.

In a further embodiment, patients that suffer from an excess of Mrg, hypersensitivity to Mrg ligands or excessive activation of Mrg may be treated by administering an effective amount of anti-sense RNA or anti-sense oligodeoxyribonucleotides corresponding to the Mrg gene coding region, thereby decreasing expression of Mrg.

As used herein, a subject to be treated can be any mammal, so long as the mammal is in need of modulation of a pathological or biological process mediated by a protein of the invention. For example, the subject may be experiencing pain or may be anticipating a painful event, such as surgery. The invention is particularly useful in the treatment of human subjects.

In the therapeutic methods of the present invention the patient is administered an effective amount of a composition of the present invention, such as an Mrg protein, peptide fragment, Mrg variant, Mrg agonist, Mrg antagonist, or anti-Mrg antibody of the invention.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular biological or pathological process. For example, an agent of the present invention can be administered in combination with other known drugs or may be combined with analgesic drugs or non-analgesic drugs used during the treatment of pain that occurs in the presence or absence of one or more other pathological processes. As used herein, two or more agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention are administered to a mammal, preferably to a human patient, in accord with known methods. Thus the agents of the present invention can be administered via parenteral, subcutaneous, intravenous, intramuscular, intraperitoneal, intracerebrospinal, intra-articular, intrasynovial, intrathecal, transdermal, topical, inhalation or buccal routes. They may be administered continuously by infusion or by bolus injection. Generally, where the disorder permits the agents should be delivered in a site-specific manner. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age,

health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

5 The toxicity and therapeutic efficacy of agents of the present invention can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. While agents that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such compounds to the desired site of action in order to reduce side effects.

10 While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. For the prevention or treatment of disease, the appropriate dosage of agent will depend on the type of disease to be treated, the severity and course of the disease, whether the agent is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the agent, and the discretion of the attending physician. Therapeutic agents are suitably administered to the patient at one time or over a series of treatments.

15 Typical dosages comprise 0.1 to 100  $\mu\text{g/kg}$  body wt. The preferred dosages comprise 0.1 to 10  $\mu\text{g/kg}$  body wt. The most preferred dosages comprise 0.1 to 1  $\mu\text{g/kg}$  body wt. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. The progress of this therapy is easily monitored by conventional techniques and assays.

20 In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition,

25 suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or

30 dextran. Optionally, the suspension may also contain stabilizers. Liposomes can also



be used to encapsulate the agent for delivery into the cell. The agent can also be prepared as a sustained-release formulation, including semipermeable matrices of solid hydrophobic polymers containing the protein. The sustained release preparation may take the form of a gel, film or capsule.

5           The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral or topical administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

10           Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

15           In practicing the methods of this invention, the compounds of this invention may be used alone or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be co-administered along with other compounds typically prescribed for these conditions according to generally accepted medical practice. The compounds of this invention can be utilized *in vivo*, ordinarily in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or *in vitro*. When used *in vivo*, the compounds must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

20

#### **a. Articles of Manufacture**

25           In another embodiment of the invention, an article of manufacture containing materials useful for the treatment of the disorders described above is provided. The article of manufacture comprises a container and a label or package insert(s) on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, etc. The containers may be formed from a variety of materials such as glass or plastic. The container holds a composition which is effective for treating the condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). At  
30           least one active agent in the composition is an Mrg agonist. The label or package insert indicates that the composition is used for treating the condition of choice, such as to

treat impaired sensory perception, for example to reduce neuropathic pain. In one embodiment, the label or package inserts indicates that the composition comprising the Mrg agonist can be used to treat pain, glaucoma or to accelerate wound healing.

## 5      **M. Transgenic Animals**

Transgenic animals containing mutant, knock-out or modified genes corresponding to the *mrg* and/or *drg-12* sequences are also included in the invention. Transgenic animals are genetically modified animals into which recombinant, exogenous or cloned genetic material has been experimentally transferred. Such genetic material is often referred to as a “transgene”. The nucleic acid sequence of the transgene, in this case a form of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 20, 22, 24, 26 or 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 7274, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106 or 108 may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene. In addition the transgene may encode a non-functional variant. The transgene may consist of nucleic acid sequences derived from the genome of the same species or of a different species than the species of the target animal.

The term “germ cell line transgenic animal” refers to a transgenic animal in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic animal to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic animals.

The alteration or genetic information may be foreign to the species of animal to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

Transgenic animals can be produced by a variety of different methods including transfection, electroporation, microinjection, gene targeting in embryonic stem cells and recombinant viral and retroviral infection (*see, e.g.*, U.S. Patent No. 4,736,866; U.S.

Patent No. 5,602,307; Mullins *et al.* Hypertension 22(4):630-633 (1993); Brenin *et al.* Surg. Oncol. 6(2):99-110 (1997); Tuan (ed.), Recombinant Gene Expression Protocols, Methods in Molecular Biology No. 62, Humana Press (1997)).

5 A number of recombinant or transgenic mice have been produced, including those which express an activated oncogene sequence (U.S. Patent No. 4,736,866); express simian SV40 T-antigen (U.S. Patent No. 5,728,915); lack the expression of interferon regulatory factor 1 (IRF-1) (U.S. Patent No. 5,731,490); exhibit dopaminergic dysfunction (U.S. Patent No. 5,723,719); express at least one human gene which participates in blood pressure control (U.S. Patent No. 5,731,489); display greater  
10 similarity to the conditions existing in naturally occurring Alzheimer's disease (U.S. Patent No. 5,720,936); have a reduced capacity to mediate cellular adhesion (U.S. Patent No. 5,602,307); possess a bovine growth hormone gene (Clutter *et al.* Genetics 143(4):1753-1760 (1996)); or, are capable of generating a fully human antibody response (McCarthy The Lancet 349(9049):405 (1997)).

15 While mice and rats remain the animals of choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species. Transgenic procedures have been successfully utilized in a variety of non-murine animals, including sheep, goats, pigs, dogs, cats, monkeys, chimpanzees, hamsters, rabbits, cows and guinea pigs (*see, e.g.*, Kim *et al.* Mol. Reprod. Dev. 46(4):  
20 515-526 (1997); Houdebine Reprod. Nutr. Dev. 35(6):609-617 (1995); Petters Reprod. Fertil. Dev. 6(5):643-645 (1994); Schnieke *et al.* Science 278(5346):2130-2133 (1997); and Amoah J. Animal Science 75(2):578-585 (1997)).

The method of introduction of nucleic acid fragments into recombination competent mammalian cells can be by any method that favors co-transformation of  
25 multiple nucleic acid molecules. Detailed procedures for producing transgenic animals are readily available to one skilled in the art, including the disclosures in U.S. Patent No. 5,489,743 and U.S. Patent No. 5,602,307.

It is contemplated that mice lacking a particular Mrg or drg-12 gene, or in which expression of a particular Mrg or drg-12 has been increased or decreased will be used in  
30 an assay for determining how Mrgs influence behavior, including sensory responses, particularly responses to painful stimuli. In particular, transgenic mice will be used to

determine if Mrg mediates the response to a particular type of noxious stimuli, such as mechanical, thermal or chemical. Thus in one embodiment transgenic mice lacking native Mrg receptors, or in which Mrg receptor expression levels have been modified, will be tested to determine their sensitivity to pressure, temperature, and other noxious stimuli. Assays for determining sensitivity to stimuli are well known in the art. These include, but are not limited to, assays that measure responsiveness to mechanical pain (von Frey hairs or tail pinch), thermal pain (latency to lick or jump in the hot plate assay), chemical pain (latency to lick when a noxious substance such as capsaicin or formalin is injected in the paw), visceral pain (abdominal stretching in response to intraperitoneal injection of acetic acid) and neuropathic pain. For example, mice in which one or more Mrgs have been deleted will be tested for their responsiveness to a variety of painful stimuli of varying intensity. By determining the sensory responses that are mediated by the Mrg receptors, therapeutic agents known to stimulate or inhibit Mrg receptors can be chosen for the treatment of disease states known to involve these types of responses. In addition, therapeutics specifically aimed at treating disorders involving these responses can be developed by targeting the Mrg receptors.

In one embodiment, transgenic mice expressing one or more human Mrg proteins are produced. The expression pattern of the human Mrg protein may then be determined and the effect of the expression of the human Mrg protein on various sensory modalities may be investigated. Further, the efficacy of potential therapeutic agents may be investigated in these mice.

In addition, the effects of changes in the expression levels of specific Mrg proteins can be investigated in animal models of disease states. By identifying the effect of increasing or decreasing Mrg receptor levels and activation, therapeutic regimens useful in treating the diseases can be developed. In one embodiment, mice in which Mrg receptor expression levels have been increased or decreased are tested in models of neuropathic pain.

Further, mice in which Mrg expression levels have been manipulated may be tested for their ability to respond to compounds known to modulate responsiveness to pain, such as analgesics. In this way the role of Mrg in the sensation of pain may be further elucidated. For example, a lack of response to a known analgesic in the

transgenic mice lacking Mrg would indicate that the Mrg receptors play a role in mediating the action of the analgesic.

Another preferred transgenic mouse is one in which the Mrg gene is modified to express a marker or tracer such as green fluorescent protein (GFP). By examining the expression pattern of the marker or tracer, the exact location and projection of Mrg containing neurons and other cells can be mapped. This information will be compared to the location and projection of neurons and other cells whose involvement in specific disease states has previously been identified. In this way additional therapeutic uses for the compounds of the present invention may be realized.

## N. Diagnostic Methods

As described in the Examples, the genes and proteins of the invention may be used to diagnose or monitor the presence or absence of sensory neurons and of biological or pathological activity in sensory neurons. For instance, expression of the genes or proteins of the invention may be used to differentiate between normal and abnormal sensory neuronal activities associated with acute pain, chronic intractable pain, or allodynia. Expression levels can also be used to differentiate between various stages or the severity of neuronal abnormalities. One means of diagnosing pathological states of sensory neurons involved in pain transmission using the nucleic acid molecules or proteins of the invention involves obtaining tissue from living subjects. These subjects may be non-human animal models of pain.

The use of molecular biological tools has become routine in forensic technology. For example, nucleic acid probes may be used to determine the expression of a nucleic acid molecule comprising all or at least part of the sequences of the invention in forensic/pathology specimens. Further, nucleic acid assays may be carried out by any means of conducting a transcriptional profiling analysis. In addition to nucleic acid analysis, forensic methods of the invention may target the proteins of the invention to determine up or down regulation of the genes (Shiverick *et al.*, Biochim Biophys Acta 393(1): 124-33 (1975)).

Methods of the invention may involve treatment of tissues with collagenases or other proteases to make the tissue amenable to cell lysis (Semenov *et al.*, Biull Eksp

Biol Med 104(7): 113-6 (1987)). Further, it is possible to obtain biopsy samples from different regions of the kidney or other tissues for analysis.

Assays to detect nucleic acid or protein molecules of the invention may be in any available format. Typical assays for nucleic acid molecules include hybridization or PCR based formats. Typical assays for the detection of proteins, polypeptides or peptides of the invention include the use of antibody probes in any available format such as *in situ* binding assays, etc. See Harlow *et al.*, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988 and Section G. In preferred embodiments, assays are carried-out with appropriate controls.

The above methods may also be used in other diagnostic protocols, including protocols and methods to detect disease states in other tissues or organs.

#### **O. Methods of Identifying Other Genes Expressed in Primary Nociceptive Sensory Neurons.**

As described in the Examples, the *mrg* and *drg-12* genes of the invention have been identified RNA using a suppression-PCR-based method (Clontech) to enrich for genes expressed in the DRG of wild type but not *Ngn1* mutant mice. This general method may be used to identify and isolate other DRG specific genes by producing transgenic mice that do not express other genes required for the development or presence of the nociceptive subset of DRG neurons. For instance, *TrkA*<sup>-/-</sup> mice may be used in the methods of the invention to isolate other genes associated with nociceptive DRG neurons (see Lindsay Philos. Trans R. Soc. Lond. B. Biol. Sci. 351(1338): 365-73 (1996) and Walsh *et al.* J. Neurosci. 19(10): 4155-68).

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

## EXAMPLES

### **Example 1: Positive selection-based differential hybridization between wild type and *Ngn1*<sup>-/-</sup> DRG to identify candidate genes involved in nociception.**

5 Previous studies have shown that Neurogenin1 (*Ngn1*), a bHLH transcription factor (Ma et al. Cell 87: 43-52 (1996)), is required for cell fate determination of nociceptive sensory neurons in dorsal root ganglia (DRG) (Ma et al. Genes & Dev. 13: 1717-1728 (1999)). In *Ngn1*<sup>-/-</sup> mutant mouse embryos most if not all *trkA*<sup>+</sup> neurons, which include the nociceptive subclass, fail to be generated. This mutant phenotype was exploited to isolate genes specifically expressed in such neurons, by subtracting  
10 cDNAs from neonatal wild-type and *Ngn1*<sup>-/-</sup> DRG. Genes expressed in the former but not the latter cDNA population are specific to *trkA*<sup>+</sup> nociceptive neurons.

Total RNA was isolated from the dorsal root ganglia (DRG) of newborn wild type or *Ngn1*<sup>-/-</sup> mice (see Ma et al. Genes Develop. 13:1717-1728 (1999), Fode et al. Neuron 20:483-494 (1998) and Ma et al. Neuron 20:469-482 (1998)). A suppression-  
15 PCR-based method (Clontech) was then used to enrich for genes expressed in wild type but not *Ngn1* mutant DRG. Briefly, cDNA was synthesized from the RNA using Superscript reverse transcriptase (Gibco) with oligo dT primers, and was amplified with the Smart PCR Amplification Kit (Clontech). The amplified wild-type and *Ngn1*<sup>-/-</sup> DRG cDNAs were used as tester and driver, respectively, in the PCR-Select subtractive  
20 hybridization protocol (Clontech). Differential screening by dot blot analysis identified several clones, which were enriched in cDNA from wild-type DRG compared to that from *Ngn1*<sup>-/-</sup> DRG. These clones were analyzed further by nucleotide sequencing and in situ hybridization.

Approximately 1,600 positives were identified in the primary screen, and of  
25 these 142 were sequenced. Fifty of these represented known genes, and 92 represented new genes (see Table 2). Among the known genes were several signaling molecules specifically expressed in nociceptive sensory neurons. These included VR-1, calcitonin gene-related peptide (CGRP), the tetrodotoxin-insensitive sodium channel (SNS-TTXi) and diacylglycerol kinase. Among the new genes were several encoding proteins with

structural features characteristic of ion channels or receptors, which were revealed by in situ hybridization to be specifically expressed in a subset of DRG sensory neurons. These molecules are described in more detail in Examples 2 and 3.

5 **Table 2. Summary of results of the differential hybridization screening for genes involved in pain sensation.**

<u># of times isolated from the screen</u>	<u>Name</u>
A. Known genes:	
13	<b>NaN</b>
9	<b>Diacylglycerol kinase</b>
7	Synaptophysin Iia
5	<b>Vanilloid receptor1</b>
3	GluR5-2c
2	<b>CGRP</b>
2	CLIM1
1	<b>SNS-TTXi</b>
1	Alpha N-catenin I
1	Brain Na channel III
1	NICA6
1	Secretogranin
B. Novel genes:	
2	Mrg3 (a novel G-protein-coupled receptor)
2	DRG12

Note: Previous studies have shown that the genes with bolded letters are expressed specifically in nociceptors.

Example 2: A novel family of putative G protein-coupled receptors  
specifically expressed in nociceptive sensory neurons.

Among the novel genes isolated from the screen were two independent clones encoding a receptor protein with 7 transmembrane segments (SEQ ID NO: 1), a characteristic of G protein-coupled receptors. The novel 7 transmembrane receptor isolated is most closely related to the oncogene *mas*, and therefore has been named *mas-related gene-3 (mrg3)*. *mrg3* is also known as *mas-related gene A1*, or *MrgA1*. A



complete coding sequence for *mrg3* has been deduced from the genomic DNA sequence (Fig. 1A and SEQ ID NO: 2). *MrgA1* shows significant homology (35% identity) to MAS1 (Young et al. Cell 45: 711-9 (1986)). It also shares significant homology (30-35% identity) with two other mammalian GPCRs, called Mas-related gene 1 (MRG1) (Monnot et al. Mol Endocrinol 5: 1477-87 (1991)) and rat thoracic aorta (RTA) (Ross et al. Proc Natl Acad Sci U S A 87: 3052-6 (1990)).

Such G protein-coupled receptors are expressed in other classes of sensory neurons, such as olfactory and gustatory neurons, but molecules in this class had not previously been described in DRG sensory neurons, with the exception of the Protease-Activated Receptors (PARs).

Further screening of mouse DRG cDNA library and mouse genomic library by using *mrg3* DNA as a probe has identified nine additional closely related genes named *mrg4* (*MrgA2*), *mrg5* (*MrgA3*), *mrg6*, *mrg7*, *mrg8* (*MrgA4*), *mrg9* (*MrgA5*), *mrg10* (*MrgA6*), *mrg11* (*MrgA7*), and *mrg12* (*MrgA8*). Among them, *mrg4*, 5 and *mrg 8-12* contain full-length open reading frames (see Fig. 1). Two human homologues were found by searching databases using the blast program. The protein alignment of the eight *mrg* genes, *mrg3-8* and human1-2, suggested that they define a novel G protein-coupled receptor gene family (Figure 1A).

In particular *MrgA1-4* were isolated from a P0 mouse DRG cDNA library and clones containing the entire ORFs of *MRGsA5-8* were isolated from a mouse genomic BAC library arrayed on filters (Incyte Genomics). Figure 6A shows an alignment of the polypeptide sequence of *MrgA1-8* and indicates the transmembrane domains as well as the cytoplasmic and extracellular loops. In addition, other mouse *MrgAs*, as well as other human *Mrg* sequences, were identified by searching the Celera mouse and human (Venter et al. Science 291: 1304-51 (2001)) genomic databases, using the TBLASTN program with *MrgA1* as the query. Table 3 shows that the *MrgA* genes are highly homologous to each other. This high degree of homology combined with the presence of certain characteristic conserved residues indicates that they define a novel subfamily of the MAS family of GPCRs.

To identify additional members of the mouse *Mrg* family, TBLASTN searches were run against the Celera mouse fragment database (indexed January 7, 2001;

18,251,375 fragments) using MRGA1 and MRGA4 protein sequences as queries. These searches identified 299 unique mouse genomic DNA fragments. The sequences of these fragments were downloaded and assembled into contigs with GELMERGE (GCG Wisconsin Package) under stringent conditions (90% identity, 20 nt minimum overlap). GELMERGE was run again (80% identity, 20 nt minimum overlap) to reduce the dataset further. The consensus nucleotide sequence from each contig was then queried against the Celera mouse fragment database with BLASTN to identify additional sequences for assembly (final n=536 fragments). The consensus sequences from the final assembly were placed into a FASTA formatted database. This database was then searched with TFASTY using MRGA1 as query to identify the potential coding regions from each consensus sequence, regardless of whether the error-prone genomic sequence introduced stop codons or frameshifts into the proteins (Pearson, W. R. (1999). Flexible similarity searching with the FASTA3 program package. In Bioinformatics Methods and Protocols, S. Misener and S. A. Krawetz, eds. (Totowa, NJ: Humana Press), pp. 185-219). The protein sequences from these searches were then combined into a single FASTA formatted file for phylogenetic analysis.

Using this analysis, 16 additional members of the murine *MrgA* subfamily were identified (Figure 6B). In addition to this subfamily, two closely related *Mrg* subfamilies called *MrgB* and *MrgC*, were also discovered (Figure 6B). To confirm the existence of an ORF in the mouse *MrgB* genes, high-fidelity PCR was used to amplify *mMrgB1-5*, *mMrgD*, and *mMrgE* from C57Bl/6 mouse genomic DNA. Several independent clones were sequenced and confirmed the ORF predictions. The presence of numerous stop codons and frame shifts in the assembled Celera sequence indicated that the *mMrgC* genes are pseudogenes.

The *MrgB* subfamily contains 14 genes, whereas *MrgC* has 12 members. The percent sequence identity within each of these subfamilies is greater than 50% (Table 3). Strikingly, all 12 *MrgC* members appear to be pseudogenes (Fig. 1B, “Ψ”), as they contain multiple premature stop codons, frameshift mutations or both. Together, therefore, the *MrgA* and *MrgB* subfamilies comprise 36 intact ORFs.

**Table 3. Similarity and identity between murine MRG subfamilies**

	mMRG A1	mMRG A2	mMRG A3	mMRG B1	mMRG B2	mMRG B3	mMRG C1	mMRG C2	mMRG C3
mMRGA1	-----	<b>77.9</b>	<b>73.1</b>	<b>48.1</b>	<b>46.3</b>	<b>43.6</b>	<b>44.9</b>	<b>46.7</b>	<b>47.8</b>
mMRGA2	87.5	-----	<b>71.8</b>	<b>42.4</b>	<b>45.4</b>	<b>42.7</b>	<b>41.5</b>	<b>44.5</b>	<b>43.5</b>
mMRGA3	85.1	83.1	-----	<b>47.9</b>	<b>46.8</b>	<b>44.2</b>	<b>46.0</b>	<b>49.8</b>	<b>46.6</b>
mMRGB1	72.1	66.8	70.2	-----	<b>57.6</b>	<b>50.0</b>	<b>42.9</b>	<b>47.1</b>	<b>45.3</b>
mMRGB2	68.7	67.7	69.4	72.7	-----	<b>53.5</b>	<b>41.8</b>	<b>44.4</b>	<b>43.1</b>
mMRGB3	65.2	65.7	64.6	69.5	73.5	-----	<b>37.0</b>	<b>38.8</b>	<b>36.4</b>
mMRGC1	69.5	65.2	70.9	64.4	67.0	63.3	-----	<b>76.0</b>	<b>79.1</b>
mMRGC2	69.8	72.5	74.2	69.4	70.8	65.7	81.4	-----	<b>78.8</b>
mMRGC3	70.9	67.2	71.0	66.2	69.5	64.6	86.1	86.3	-----

Percent identity (top-right, bold) and percent similarity (bottom-left) between the protein sequences are indicated. “hMRG” indicates a human MRG amino acid sequence; “mMRG” indicates a murine MRG sequence. “hMRGX” is used to indicate a human homolog of mMRGA and mMRGB sequences (Fig. 1B). Values were derived from global alignments using the GAP program in the GCG package.

Searches of the Celera (Venter et al. Science 291: 1304-51 (2001)) and public (Consortium. Nature 409: 860-921 (2001)) genomic sequence databases, using both BLAST (Altschul et al. Journal of Molecular Biology 215: 403-410 (1990)) and Hidden Markov Models (HMMs (Eddy. Bioinformatics 14, 755-63 (1998))), revealed 4 closely related (~50% identity) full-length human genes, and at least 10 human pseudogenes. Briefly, TBLASTN searches were run against the Celera human genome database (Venter et al. Science 291: 1304-51 (2001)) using the *mMrgA1* protein sequence as the query. The genomic sequences that were identified in this search were downloaded, placed into a FASTA formatted database and searched with TFASTY to identify a non-redundant set of proteins. With the exception of *hMrgX3*, *hMrgE*, and *hMrgΨ8*, all human *Mrgs* were independently identified from a similar analysis of the public human genome sequence (Consortium. Nature 409: 860-921 (2001)). Human *MrgX1-4* sequences were independently verified from PCR-amplified products derived from human BAC clones containing the genes.

Although the human genes appear to be more similar to the murine *MrgA* subfamily than the *MrgB* subfamily in the phylogenetic tree (Fig. 6B, hMrgX1-4), in the

absence of clear orthologous pairs we currently refer to them as *hMrgX* genes. In addition to the *MrgA*, *B* and *C* subfamilies, a number of additional *Mas1*-related orphan GPCRs were identified by this search, including those we refer to as *Mrgs D-F* (Fig. 6B). Several of these sequences, such as *MrgD*, have clear human orthologs (Fig. 6B, *hMrgD* and Table 4). All together, we identified almost 45 murine and 9 human intact coding sequences belonging to this family.

**Table 4. Similarity and identity between human and murine MRGs**

	hMRG X2	hMRG D	hMRG E	mMRG A1	mMRG B4	mMRG B1	mMRG D	mMRG E
hMRGX2	-----	<b>39.3</b>	<b>40.2</b>	<b>55.6</b>	<b>50.1</b>	<b>53.4</b>	<b>40.5</b>	<b>38.8</b>
hMRGD	65.4	-----	<b>34.4</b>	<b>37.6</b>	<b>35.4</b>	<b>33.8</b>	<b>55.8</b>	<b>35.9</b>
hMRGE	62.8	54.6	-----	<b>36.6</b>	<b>32.8</b>	<b>32.8</b>	<b>33.9</b>	<b>76.5</b>
mMRGA1	74.8	63.4	57.7	-----	<b>48.1</b>	<b>48.1</b>	<b>37.1</b>	<b>39.7</b>
mMRGB4	71.0	64.0	58.0	70.4	-----	<b>54.5</b>	<b>34.8</b>	<b>36.6</b>
mMRGB1	73.5	58.6	60.5	72.1	74.1	-----	<b>36.5</b>	<b>33.8</b>
mMRGD	61.1	72.6	57.6	59.5	64.2	61.3	-----	<b>35.1</b>
mMRGE	59.0	59.5	84.0	62.5	63.7	59.1	59.3	-----

Percent identity (top-right, bold) and percent similarity (bottom-left) between the protein sequences are indicated. "hMRG" indicates a human MRG amino acid sequence; "mMRG" indicates a murine MRG sequence. "hMRGX" is used to indicate a human homolog of mMRGA and mMRGB sequences (Fig. 1B). Values were derived from global alignments using the GAP program in the GCG package.

MRG receptors have short (3-21 amino acid) N-termini with no apparent signal peptide, which are predicted to be located extracellularly. The transmembrane domains and intracellular domains are highly conserved suggesting that the receptors have a shared function. The most divergent regions of MRGA-family receptors appear localized to the extracellular loops (Fig. 6A), suggesting that these receptors recognize different ligands, or the same ligand but with different affinities. Interestingly, we identified 12 single nucleotide polymorphisms in the *MrgA1* coding sequence between murine strains C57BL/6J and 129SvJ. These 12 changes resulted in 6 amino acid

substitutions, all of which were either conservative, or which substituted residues expressed at the same position by other family members.

A large mouse genomic contig was built by analyzing overlapping BAC clones containing *MrgA* sequences (Fig. 6C). There are 7 *MrgA* genes, including 3 pseudogenes, residing in this contig. Such clustering is a common feature of GPCR-encoding gene families (Xie et al. Mamm Genome 11: 1070-8 (2000)). Strikingly, all of the human *Mrg* genes (with the exception of *Mas1* and *Mrg1*) are located on chromosome 11, which also contains 50% of all human olfactory receptors genes. All of the *MrgA* genes in the murine BAC contig (Fig. 6C) encode intact ORFs with N-terminal methionines, like many other GPCR-encoding genes. Using the Celera mouse genome database, sequences flanking each *MrgA* coding region were obtained and analyzed. This analysis revealed that at least six *MrgA* genes have L1 retrotransposon sequences located ~650 bp downstream of their coding sequences (Fig. 6B, indicated by “L1”).

All of the eight full-length *mas*-related genes, *mrg3-5* and *mrg8-12*, are enriched in nociceptive sensory neurons as indicated by their expression in a subset of DRG sensory neurons which are eliminated in *ngn1*<sup>-/-</sup> mutant DRG (Fig 2 and 2A).

### Example 3: A novel two-transmembrane segment protein specifically expressed in nociceptive sensory neurons.

Another novel gene isolated in this screen, *drg12* (SEQ ID NO: 13), encodes a protein with two putative transmembrane segments (SEQ ID NO: 14). In situ hybridization indicates that, like the *mrg* genes, this gene is also specifically expressed in a subset of DRG sensory neurons. Although there are no obvious homologies between this protein and other sequences in the database, it is noteworthy that two purinergic receptors specifically expressed in nociceptive sensory neurons (P<sub>2</sub>X<sub>2</sub> and P<sub>2</sub>X<sub>3</sub>) have a similar bipartite transmembrane topology. Therefore it is likely that *drg12* also encodes a receptor or ion channel involved in nociceptive sensory transduction or its modulation. The hydrophobicity of a homologous region of a *drg12* human sequence (SEQ ID NO: 19) is compared with the hydrophobicity of mouse *drg12* in Fig. 4.

Example 4: *mrg* and *drg-12* genes are  
specifically expressed in nociceptive sensory neurons.

The prediction of function for *mrg*-family and *drg-12* genes is based on their structure and expression pattern, taken together with the identification of ligands as described below. To determine whether Mrg proteins are expressed in DRG neurons, in situ hybridization using dioxxygenin-labeled riboprobes was performed. Briefly, tissue was obtained from P0 mouse pups and fixed in 4% paraformaldehyde overnight at 4°C, cryoprotected in 30% sucrose overnight and embedded in OCT. Tissue sections were cut transversely on a cryostat at 18 µm. Non-isotopic in situ hybridization on frozen sections was performed as previously described using cRNA probes (Ma et al. Cell 87: 43-52 (1996); Perez et al. Development 126: 1715-1728 (1999)). Eight *MrgAs*, 5 *MrgBs* and *MrgD* were used as probes. At least 10 DRGs were analyzed to count the number of neurons positive for each probe.

*Mrg* and *drg12* genes, including all eight *MrgAs* (*MrgA1-8*), are expressed in subsets of small-diameter sensory neurons in the dorsal root ganglia (DRG) of the mouse (Fig. 7B-I). Importantly, the expression of all eight *MrgAs* was virtually absent in the DRGs of *Ngn1*<sup>-/-</sup> animals (Figure 7J), consistent with the design of the subtractive hybridization screen. Among the eight *MrgA* clones examined, *MrgA1* has the widest expression within sensory neurons in DRGs (13.5%). Other *MrgAs* are only expressed in several cells per DRG section (ranging from 0.2-1.5% of DRG neurons). This differential abundance may explain why only *MrgA1* was isolated in the original screen. No obvious differences in the expression patterns of *MrgA1-8* were noticed in DRGs from different axial levels. This expression is highly specific, in that expression of these genes has thus far not been detected in any other tissue of the body or in any other region of the nervous system thus far examined.

Like the *MrgA* genes, *MrgD* was also specifically expressed in a subset of DRG sensory neurons (see below, Figure 15). In contrast, *MrgB1-5* were not detectably expressed in DRGs. However, mMrgB1 expression has been observed in scattered cells in the epidermal layer of skin in newborn mice, as well as in the spleen and the

submandibular gland (Figures 13 and 14). These cells appear to be immune cells that play a role in wound repair. mMrgB2 also shows this expression pattern. In contrast, mMrgB3, mMrgB4 and mMrgB5 do not appear to be expressed in any of these tissues.

These results indicate that *Mrg* and *drg12* genes are expressed in primary sensory neurons. However, DRG contain different classes of neurons subserving different types of sensation: e.g., heat, pain, touch and body position. Independent identification is provided by the fact that the neurons that express the *mrg*-family and *drg12* genes are largely or completely eliminated in *Ngn1*<sup>-/-</sup> DRG (Figure 2), because the *Ngn1* mutation is independently known to largely or completely eliminate the nociceptive (noxious stimuli-sensing) subset of DRG neurons, identified by expression of the independent markers *trkA*, *VR-1* and *SNS-TTXi* (Ma *et. al.* Genes & Dev. 13: 1717-1728 (1999)). The loss of *mrg*- and *drg12*- expressing neurons in *Ngn1*<sup>-/-</sup> mutant DRG therefore indicates that these genes are very likely expressed in nociceptive sensory neurons. Although small numbers of sensory neurons of other classes (*trkB*<sup>+</sup> and *trkC*<sup>+</sup>) are eliminated in the *Ngn1*<sup>-/-</sup> mutant as well, *mrg* and *drg12* genes are unlikely to be expressed in these classes of sensory neurons, because if they were then the majority of *mrg*- and *drg12*-expressing sensory neurons would be predicted to be spared in the *Ngn1*<sup>-/-</sup> mutant, and that is not the case.

The lack of expression of *MrgAs* in DRGs from *Ngn1*<sup>-/-</sup> mice is consistent with the idea that they are expressed in cutaneous sensory neurons. Furthermore, the distribution of *MrgA1*<sup>+</sup> cells was similar to that of neurons expressing *trkA*, a marker of nociceptive sensory neurons (McMahon *et al.* Neuron 12: 1161-71 (1994); Snider and Silos-Santiago Philos Trans R Soc Lond B Biol Sci 351: 395-403 (1996)) (Fig. 7A, B). To directly determine whether *MrgA* genes are expressed in *trkA*<sup>+</sup> cells, in situ hybridization was performed for *MrgA1*, *A3* and *A4* in conjunction with immunolabeling using anti-*trkA* antibodies, on neonatal DRG. Fluorescein-UTP-labeled cRNA probes were detected with alkaline phosphatase- (AP-) conjugated anti-fluorescein antibody (1:2000, Roche) and developed with Fast Red (Roche) to generate a red fluorescent signal. After the fluorescent in situ hybridization was performed, sections were incubated in primary antibodies against *TrkA* (1:5000, gift from Dr. Louis Reichardt), *VR1* (1:5000, gift from Dr. D. Julius), *CGRP* (1:500, Chemicon), or

SubstanceP (1:1000, Diasorin). All antibodies were diluted in 1x PBS containing 1% normal goat serum and 0.1% TritonX-100. Primary antibody incubations were carried out overnight at 4 °C. Secondary antibodies used were goat-anti-rabbit-IgG conjugated to Alexa 488 (1:250, Molecular Probes). For double-labeling with *Griffonia simplicifolia* IB4 lectin, sections were incubated with 12.5 µg/ml FITC-conjugated IB4 lectin (Sigma) following in situ hybridization.

Double labeling experiment using *mrgs* antisense RNA probes with anti-trkA antibodies confirmed that *mrgs*, specifically *MrgAs*, are co-expressed by trkA<sup>+</sup> nociceptive neurons in DRG (see Fig. 7B and Fig. 8A-C). Similar results were obtained for *MrgD* (Fig. 8D). Taken together, these data indicate that *MrgAs* and *MrgD* are specifically expressed by nociceptive sensory neurons in DRG.

Further experiments were carried out to determine whether *Mrgs* are expressed in particular subsets of nociceptors. Additional double labeling experiments using *mrgs* antisense RNA probes with anit-VR1 and isolectin B4 (IB4)-labeling, as described above, have shown that *mrgs* are preferentially expressed by IB4<sup>+</sup> nociceptive neurons but not VR1-expressing nociceptive neurons (Fig. 2C and 2D). In particular, combined fluorescent labeling for IB4 together with in situ hybridization with *MrgA1*, *A3*, *A4* and *MrgD* probes clearly showed that these receptors are expressed by IB4<sup>+</sup> neurons (Fig. 8E-H), and may be restricted to this subset. This result indicates that these *Mrgs* are expressed by non-peptidergic nociceptive neurons that project to lamina IIi (Snider and McMahon Neuron 20: 629-32 (1998)). Consistent with this assignment, the majority (90%) of *MrgA1*<sup>+</sup>, and all *MrgA3*<sup>+</sup>, *A4*<sup>+</sup> and *MrgD*<sup>+</sup> cells, lack substance P expression (Fig. 8I-L). Similarly, the majority (70%) of *MrgA1*<sup>+</sup>, and all *MrgA3*<sup>+</sup>, *A4*<sup>+</sup> and *MrgD*<sup>+</sup> cells, do not express CGRP (Fig. 8M-P), another neuropeptide expressed by C-fiber nociceptors. Previous studies had shown that IB4<sup>+</sup> nociceptive neurons were involved in neuropathic pain resulting from nerve injury (Malmberg, A. B. et al. Science 278: 279-83 (1997)). Neuropathic pain including postherpetic neuralgia, reflex sympathetic dystrophy, and phantom limb pain is the most difficult pain to be managed. *Mrgs* may play essential roles in mediating neuropathic pain and may provide alternative solutions to manage neuropathic pain.



Recent studies have provided evidence for the existence of two neurochemically and functionally distinct subpopulations of IB4<sup>+</sup> nociceptors: those that express the vanilloid receptor VR1 (Caterina et al. Science 288: 306-13 (1997)), and those that do not (Michael and Priestley J Neurosci 19: 1844-54 (1999); Stucky and Lewin J Neurosci 19: 6497-505 (1999)). Strikingly, in situ hybridization with *MrgA* or *D* probes combined with anti-VR1 antibody immunostaining indicated that the *MrgA1*, *A3*, *A4* and *D*-expressing cell population was mutually exclusive with VR1<sup>+</sup> cells (Fig. 8Q-T). In summary, these expression data demonstrate that *MrgA* and *D* genes are expressed in the subclass of nonpeptidergic cutaneous sensory neurons that are IB4<sup>+</sup> and VR1<sup>-</sup> (Fig. 9).

#### *MrgA1* is co-expressed with other *MrgA* genes

*MrgA1* is more broadly expressed than are the other *MrgA* genes (Fig. 2), suggesting *MrgA1* and *MrgA2-8* are expressed by different or overlapping subsets of nociceptors. Double-label in situ hybridization studies using probes labeled with digoxigenin and fluorescein indicated that most or all neurons expressing *MrgA3* or *MrgA4* co-express *MrgA1* (Fig. 10A-F). Interestingly, the fluorescent in situ hybridization signals for *MrgA3* and *A4* using tyramide amplification often appeared as dots within nuclei that were circumscribed by the cytoplasmic expression of *MrgA1* mRNA, detected by Fast Red (Fig. 10F). Such dots were not observed using the less-sensitive Fast Red detection method, and were only observed in the nuclei of *MrgA1*<sup>+</sup> cells. Similar intranuclear dots have previously been observed in studies of pheromone receptor gene expression, and have been suggested to represent sites of transcription (Pantages and Dulac Neuron 28: 835-845 (2000)). The results for *MrgA1*, 3 and 4 indicate that those neurons that express the rarer *MrgA* genes (*MrgA2-8*) are a subset of those that express *MrgA1*.

To address the question of whether *MrgA2-8* are expressed in the same or in different neurons, the number of neurons labeled by single probes was compared to that labeled by a mixture of all 7 probes (Buck and Axel Cell 65: 175-187 (1991)). Approximately 3-fold more neurons (4.5% vs. 1%) were labeled by the mixed probe than by an individual probe to *MrgA4* (Fig. 10J, K), indicating that these genes are not

all co-expressed in the same population of neurons. However, the percentage of neurons labeled by the mixed probe (4.5%) was less than the sum of the percentage of neurons labeled by each of the 7 individual probes (6.6%), indicating that there is some overlap in the expression of *MrgA2-A8*. In addition, higher signal intensity was observed in individual neurons using the mixed probe, than using a single probe.

Double-labeling experiments with *MrgA1* and *MrgD* probes were also performed. These proteins share only 60% sequence similarity, as shown in Fig. 6B and Table 3. The results of these experiments indicated only partial overlap between neurons expressing these two receptors (Fig. 10G-I). Approximately 15% (118/786) of neurons expressing either *MrgA1* or *MrgD* co-expressed both genes. Thirty-four percent (118/344) of *MrgA1*<sup>+</sup> cells co-expressed *MrgD*, while 26.7% (118/442) of *MrgD*<sup>+</sup> cells co-expressed *MrgA1*.

Taken together, these data indicate the existence of at least three distinct subpopulations of IB4<sup>+</sup>, VR1<sup>-</sup> sensory neurons: *MrgA1*<sup>+</sup>*MrgD*<sup>+</sup>; *MrgA1*<sup>+</sup>*MrgD*<sup>-</sup> and *MrgA1*<sup>-</sup>*MrgD*<sup>+</sup>. The *MrgA1*<sup>+</sup> subset is further subdivided into different subsets expressing one or more of the *MrgsA2-A8*.

#### *Mrg*-family genes encode putative G-protein coupled receptors (GPCRs).

Hydrophobicity plots of the encoded amino acid sequences of the *mrg*-family genes predicts membrane proteins with 7 transmembrane segments. Such a structure is characteristic of receptors that signal through “G-proteins.” G proteins are a family of cytoplasmic molecules that activate or inhibit enzymes involved in the generation or degradation of “second messenger” molecules, such as cyclic nucleotides (cAMP, cGMP), IP<sub>3</sub> and intracellular free calcium (Ca<sup>++</sup>). Such second messenger molecules then activate or inhibit other molecules involved in intercellular signaling, such as ion channels and other receptors.

G protein-coupled receptors (GPCRs) constitute one of the largest super-families of membrane receptors, and contain many subfamilies of receptors specific for different ligands. These ligands include neurotransmitters and neuropeptides manufactured by the body (e.g., noradrenaline, adrenaline, dopamine; and substance P, somatostatin,

respectively), as well as sensory molecules present in the external world (odorants, tastants).

Although the *mrg*-family genes are highly homologous, the most divergent regions were the extracellular domains (see Figure 6A). The variability of the extracellular domains of *mrg* family suggests that they may recognize different ligands.

The fact that the *mrg*-family genes encode GPCRs, and are specifically expressed in nociceptive sensory neurons, suggest that these receptors are involved, directly or indirectly, in the sensation or modulation of pain, heat or other noxious stimuli. Therefore the *mrg*-encoded receptors are useful as targets for identifying drugs that effect the sensation or modulation of pain, heat or other noxious stimuli. The nature of the most useful type of drug (agonistic or antagonistic) will reflect the nature of the normal influence of these receptors on the sensation of such noxious stimuli. For example, if *mrg*-encoded receptors normally act negatively, to inhibit or suppress pain, then agonistic drugs would provide useful therapeutics; conversely, if the receptors normally act positively, to promote or enhance pain, then antagonistic drugs would provide useful therapeutics. There might even be certain clinical settings in which it would be useful to *enhance* sensitivity to noxious stimuli, for example in peripheral sensory neuropathies associated with diabetes.

The nature of the influence of *mrg*-encoded GPCRs on pain sensation may be revealed by the phenotypic consequences of targeted mutation of these genes in mice. For example, if such mice displayed enhanced sensitivity to noxious stimuli, then it could be concluded that the receptors normally function to inhibit or suppress pain responses, and vice-versa. Alternatively, high-throughput screens may be used to identify small molecules that bind tightly to the *mrg*-encoded receptors. Such molecules would be expected to fall into two categories: agonists and antagonists. Agonists would be identified by their ability to activate intracellular second messenger pathways in a receptor-dependent manner, while antagonists would inhibit them. Testing of such drugs in animal models of pain sensitivity will then reveal further information concerning the function of the GPCRs: for example, if the molecules behave as receptor antagonists *in vitro*, and they suppress sensitivity or responsiveness to noxious stimuli *in vivo*, then it may be concluded that the receptor normally functions

to promote or enhance pain sensation. Conversely, if receptor agonists suppress, while antagonists enhance, pain sensation *in vivo*, then it may be concluded that the receptor normally functions to suppress or inhibit pain sensation.

#### *drg12* encodes a putative transmembrane signaling molecule

5 Hydrophobicity plots of the encoded amino acid sequence of the *drg12* gene predicts a membrane protein with 2 transmembrane segments. The membrane localization of this protein has been verified by immuno-staining of cultured cells transfected with an epitope-tagged version of the polypeptide. Although the DRG12 amino acid sequence has no homology to known families of proteins, its bipartite  
10 transmembrane structure strongly suggests that it is involved in some aspect of intercellular signaling, for example as a receptor, ion channel or modulator of another receptor or ion channel. This prediction is supported by the precedent that two known receptors with a similar bipartite transmembrane topology, the purinergic  $P_2X_2$  and  $P_2X_3$  receptors, are like DRG12, specifically expressed in nociceptive sensory neurons.

15 Based on this structural data, and its specific expression in nociceptive sensory neurons, it is probable that DRG12 is involved, directly or indirectly, in the sensation or modulation of noxious stimuli. Accordingly, the *drg12*-encoded protein is a useful target for the development of novel therapeutics for the treatment of pain.

#### 20 Example 5: *Mrg* proteins are receptors for neuropeptides.

As discussed above, the structure of the proteins encoded by *Mrg* genes indicates that they function as receptors. To identify ligands for the *Mrg* receptors, selected *MrgA* genes were tested in a calcium release assay. *MrgA* genes, including *MrgA1* and *MrgA4*, were cloned into a eukaryotic expression vector and transfected into  
25 human embryonic kidney (HEK) 293 cells. HEK-293 cells were obtained from the ATCC and cultured in DMEM supplemented with 10% fetal bovine serum. An HEK293-  $G\alpha_{15}$  cell line stably expressing  $G\alpha_{15}$  was provided by Aurora Biosciences Corporation and grown on Matrigel™ (growth factor reduced Matrigel, Becton Dickinson, diluted 1 : 200 with serum-free DMEM)-coated flasks and maintained at 37

°C in DMEM (GibcoBRL) supplemented with 10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 25 mM HEPES and 3 µg/ml blasticidin-S. For transfection, cells were seeded on Matrigel-coated 35 mm glass-bottom dishes (Biotech Inc., Butler, PA). After 16-24 hr, cells were transfected using FuGENE 6 (Roche). Transfection efficiencies were estimated by visualization of GFP fused to the C-terminus of MrgA1 and A4, and were typically > 60 %. Fusing GFP to the C-termini of the *MrgA* coding sequences additionally allowed for visual confirmation of the intracellular distribution of the receptors and their membrane integration in the transfected cells (Fig. 11D).

To increase the sensitivity of the calcium release assay, in some experiments the MRGA-GFP fusion proteins were expressed in HEK 293 cells modified to express  $G\alpha_{15}$ , which couples GPCRs to a signal transduction pathway leading to the release of intracellular free  $Ca^{2+}$  (Offermann and Simon J Biol Chem 270: 15175-80 (1995)). This calcium release can be monitored ratiometrically using Fura-2 as a fluorescent indicator dye (Tsien et al. Cell Calcium 6: 145-57 (1985)) (Fig. 11A-C). This heterologous expression system has been previously used to identify ligands for taste receptors (Chandrashekar et al. Cell 100: 703-11 (2000)).

Because MRGAs exhibit the highest sequence similarity to peptide hormone receptors, approximately 45 candidate peptides were screened for their ability to activate MRGA1, using the intracellular  $Ca^{2+}$ -release assay. Briefly, transfected cells were washed once in Hank's balanced salt solution with 11 mM D-glucose and 10 mM HEPES, pH 7.4 (assay buffer) and loaded with 2 µM Fura-2 AM (Molecular Probes) at room temperature for 90 min, with rotation. Loaded cells were washed twice with assay buffer and placed on a micro-perfusion chamber (Biotech). The chamber was mounted on top of a Olympus IMT2 inverted microscope, and imaged with an Olympus DPlanApo 40X oil immersion objective lens. Samples were illuminated by a 75W xenon bulb, and a computer-controlled filter changer (Lambda-10; Shutter Instruments) was used to switch the excitation wavelength. A cooled CCD camera (Photometric) was used in detecting fluorescence. GFP-positive cells within a field were identified using an excitation wavelength of 400 nm, a dichroic 505 nm long-pass filter and an emitter bandpass of 535 nm (Chroma Technology). In the same field, calcium

measurements were performed at an excitation wavelength of 340 nm and 380 nm, and an emission wavelength of 510 nm. Agonists were diluted in assay buffer and solution changes accomplished by micro-perfusion pump (Biotech). Fura-2 fluorescence signals (340 nm, 380 nm and the 340/380 ratio) originating from GFP-positive cells were continuously monitored at 0.4- or 1-second intervals and collected using Axon Imaging Workbench 4.0 software (Axon). Instrument calibration was carried out with standard calcium solutions (Molecular probes) in glass bottom dishes (MatTek Corp.).

At a concentration of 1  $\mu$ M, numerous neuropeptides produced some level of activation of *MrgA1*-expressing cells (Fig. 12A). These included ACTH, CGRP-I and – II, NPY and somatostatin (SST). Nevertheless, many other peptide hormones did not activate MRGA1, including angiotensins I-III and neurokinins A and B, alpha-MSH and gamma2-MSH (Fig. 12A and data not shown). *MrgA1* was only very weakly activated by eicosanoid ligands such as Prostaglandin-E1 and Arachidonic Acid (data not shown).

The most efficient responses in *MrgA1*-expressing HEK cells were elicited by RFamide peptides, including FLRF and the molluscan cardioactive neuropeptide FMRFamide (Price and Greenberg Science 197: 670-671 (1977)) (Phe-Met-Arg-Phe-amide) (Fig. 11C, 12A). Two mammalian RFamide peptides, NPAF and NPFF, which are cleaved from a common pro-peptide precursor (Vilim et al. Mol Pharmacol 55: 804-11 (1999)) were then tested. The response of *MrgA1*-expressing cells to NPFF at 1  $\mu$ M was similar to that seen with FMRFamide, while that to NPAF was significantly lower (Fig. 12A). *MrgA1* was also weakly activated by two other RFamide ligands,  $\gamma$ -MSH and schistoFLRF (data not shown).

In order to examine further the specificity of activation of *MrgA1* and *A4*, the top candidate ligands emerging from the initial screen were tested on these same receptors expressed in HEK cells lacking  $G\alpha_{15}$ . *MrgA1* and *A4* expressed in this system retained responses to RFamide peptides (Fig. 12B, C), demonstrating that the intracellular  $Ca^{2+}$  release responses seen in the initial screen are not dependent on the presence of exogenous  $G\alpha_{15}$ . This indicates that MrgAs act in HEK cells via Gq or Gi. The response of *MrgA1*-expressing HEK cells to NPFF was lower than that to FLRF (Fig. 12B), and there was no response to NPAF. Conversely, *MrgA4*-expressing cells responded to NPAF, but not to NPFF or FLRF (Fig. 12C). In both cases, the response

to NPY seen in  $G\alpha_{15}$ -expressing cells (Fig. 11A) was lost completely, while those to CGRP-II and ACTH were considerably diminished.

In order to determine the lowest concentrations of RFamide ligands capable of activating MrgA1 and A4, dose-response experiments were carried out in HEK cells expressing  $G\alpha_{15}$ , which afforded greater sensitivity (Fig. 12D, E). These experiments indicated that *MrgA1* could be activated by FLRF at nanomolar concentrations (Fig. 12D;  $EC_{50} \approx 20$  nM), and by NPFF at about an order of magnitude higher concentration (Fig. 12D;  $EC_{50} \approx 200$  nM), whereas NPAF was much less effective. In contrast, *MrgA4* was well activated by NPAF (Fig. 12E;  $EC_{50} \approx 60$  nM), and much more weakly activated by FLRF and NPFF. Neither receptor showed strong activation in response to RFRP-1, -2 or -3, a series of RFamide ligands produced from a different precursor (Hinuma et al. Nat Cell Biol 2: 703-8 (2000)). These data confirm that *MrgA1* and *MrgA4* display different selectivities towards different RFamide ligands in this system. By contrast, these receptors responded similarly to ACTH ( $EC_{50} \sim 60$ - and 200 nM for *MrgA1* and *A4*, respectively; data not shown).

Finally, given the sequence similarity between MRGA receptors and MAS1, the responsiveness of cells expressing exogenous *Mas1* to NPFF, NPAF and FLRF was tested. MAS1 showed a profile distinct from both *MrgA1* and *MrgA4* (Fig. 12F): like *MrgA1*, it was activated by NPFF at a similar concentration of the peptide ( $EC_{50} \approx 400$  nM), but unlike *MrgA1* it was poorly activated by FLRF. In contrast to *MrgA4*, MAS1 did not respond well to NPAF. No response was detected in MAS1-expressing cells upon exposure to Angiotensins I and II, ligands which have been previously reported to activate this receptor (Jackson, T. R., et al. Nature 335: 437-40 (1988)). Nor did MAS1 respond to ACTH. Thus, MAS1, *MrgA1* and *MrgA4* expressed in this heterologous system are all activated by RFamide family ligands, but with differing ligand-sensitivities and -selectivities (Table 4).

**Table 4. Selectivity of activation of Mas-related GPCRs by RF-amide ligands in HEK cells**

receptor	A. <u>Ligand</u>		
	FLRF	NPFF	NPAF
MRGA1	+++	++	+/-
MRGA4	+/-	+/-	+++
MAS1	+/-	++	+/-

5                   Relative efficacy of activation of the indicated receptors by the indicated ligands is shown. For quantification, see Fig. 6. “+++” indicates  $10 \text{ nM} < EC_{50} < 100 \text{ nM}$ ; “++” indicates  $100 \text{ nM} < EC_{50} < 500 \text{ nM}$ ; “+/-” indicates weak response seen at  $1 \text{ }\mu\text{M}$ . For details see Fig. 6.

10                   A novel family consisting of close to 50 MAS1 related g-protein coupled receptors has been identified. The specific expression of several classes of these receptors in a subset of nociceptive sensory neurons indicates that these receptors play a role in the sensation or modulation of pain. Consistently, these receptors have been shown to be activated by RFamide neuropeptides, which are known to mediate

15                   analgesia. As a result, these receptors provide a novel target for anti-nociceptive drugs.

                    Although the present invention has been described in detail with reference to examples above, it is understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by

20                   the following claims. All cited patents, patent applications and publications referred to in this application are herein incorporated by reference in their entirety.